LINEAR FILTERS ON A SEQUENCE SPACE

By

T. F. BRIDGLAND, JR.

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL OF
THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA
June, 1959

To my father and mother

ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to--and appreciation of--Dr. F. W. Kokomoor, Head of the Department of Mathematics and Chairman of the author's Supervisory Committee, without whose friendship, encouragement and assistance at several junctures the author would have found the successful completion of his graduate study in mathematics immeasurably more difficult, if not impossible; Dr. R. C. Meacham, Co-chairman of the author's Supervisory Committee, under the direction of whom the research leading to this dissertation was carried out and whose stimulating guidance, incisive criticisms and helpful comments were of inestimable value in its preparation; Dr. W. A. Nash, member of the author's Supervisory Committee, who, as director of a research group--of which the author was a member--in the Department of Engineering Mechanics, made it possible for the author to devote a portion of his work day for several months to the research reported in this dissertation; Drs. C. B. Smith, R. W. Cowan and J. T. Moore, members of the author's Supervisory Committee, for their tutelage throughout the author's graduate study and for their helpful criticism of the present study; Helen, the author's wife, whose devotion, encouragement and good humor smoothed so many rough spots and who--more than anyone else--made possible the execution of this work; Michael, Ryan and Jeffery, the author's very young sons, who waited so patiently for this work to be completed.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

INTRODUCTION 1

Chapter

I. INFINITE MATRICES AND LINEAR FILTERS ON σ 4

- Introduction
- Definitions
- Algebraic character of $\mathcal{L}$ and $\mathcal{D}$. Reciprocals
- Convergence and Summability in σ, $\mathcal{L}$ and $\mathcal{D}$
- Displacement transformations on σ

II. FORMULATION OF THE THEORY OF SAMPLED-DATA CONTROL 20

- Introduction
- Sampled-data systems as linear filters on σ
- Explicit form of constant linear filters
- Error-sampled feedback systems
- Remarks

III. NUMERICAL INTEGRATING OPERATORS 31

- Introduction
- Elementary considerations
- Type I operators
- Type II operators
- Type IV operators
- Bounds on the error for operators of types I, II and IV
- Approximate solution of linear ordinary differential equations
- Accuracy of the approximate solution
- The solution of (3.69)
- Remarks

COMPENDIUM 53

LIST OF REFERENCES 56

APPENDIX 60

- Definition of "algebra"
- Derivation of relations reciprocal to those of (3.54)
- Proof of boundedness of expressions contained in (3.72)

BIOGRAPHICAL SKETCH 65

INTRODUCTION

The operational methods of Heaviside--as embodied in a consistent application of the Laplace transform--have made a profound impression upon several areas of applied mathematics and engineering. Indeed, it is safe to say that at least one field--the theory and practice of automatic control systems--probably would not have survived its infancy without the nursemaid of operational calculus. Thus we find that the mathematical language of the contemporary control theorist or engineer is couched almost exclusively in terms of the complex frequency plane, i.e., the domain of the Laplace transform variable.

During World War II, a new class of control systems aroused the interest of control analysts. These systems--known generically as sampled-data systems--differed from control systems theretofore considered in that the input signal or actuating error of the new type of system was known, or could be specified, only on a denumerable subset of the temporal continuum. Thus, an acceptable mathematical model of such a signal can be taken to be an infinite sequence of real numbers.

Since, as has been indicated, operational methods are the *sine qua non* of control system analysis, it was natural that control analysts should attempt the application of similar methods to the study of the new type of system. However, inasmuch as the workers in this new field evidently were acquainted with the theory of linear operators only insofar as it is exemplified by use of the Laplace transform, the attempted analysis of sampled-data systems involved, fundamentally, a modification or variant of this transform. Of the success attending these attempts more will be said shortly.

At about the same time that interest was being directed toward the new class of control systems, investigations were initiated concerning the development of an operational technique for the approximate--or numerical--analysis of the older class of control systems which have come to be known as continuous-data systems to distinguish them from sampled-data systems. Here again the mathematical entity of fundamental importance was--and is--the infinite sequence of real numbers and, again, operational methods involving modifications of the Laplace transform have played a dominant role in these investigations.

The developments, on the one hand, of the theory of sampled-data systems and, on the other, of the just-mentioned approximate methods have proceeded independently, albeit concurrently. In both instances, application of the Laplace-operator methods has been an outstanding success from a purely practical engineering standpoint. However, what has been developed purports to be a mathematical theory and it is on a mathematical basis that this theory is open to considerable criticism. We shall not labor this criticism here but will comment to a greater extent upon it in the introductory remarks to Chapters II and III. It suffices for the present to state that the developments we have been discussing cannot be considered mathematically rigorous.

Certainly one function of the applied mathematician is to endeavor where possible to render rigorous--to paraphrase Bellman--the half-intuitive but wholly ingenious mathematical techniques of his colleagues in the physical sciences. When successful, such endeavors--in order to avoid the stigma of triviality--should lead to new conclusions about the systems with which they are concerned or at least open a path to the discovery of new results.

These remarks lead up to the purpose of this dissertation, which is to present a rigorous and, it is hoped, non-trivial formulation of the mathematical theory of sampled-data systems and of the approximate operational technique discussed earlier. This purpose is accomplished in the following way: in Chapter I, portions of the theory of infinite matrices are employed in the study of certain algebras of linear transformations on the space of all sequences with complex entries. In Chapter II, the considerations of the first chapter are employed to formulate a mathematical theory of sampled-data systems. Finally, in Chapter III, the same considerations are utilized in the derivation of an operational technique for numerical quadrature, this technique in turn being applied to the approximate solution of linear ordinary differential equations.

CHAPTER I

INFINITE MATRICES AND LINEAR FILTERS ON σ

Introduction.--In this chapter, two special classes of infinite matrices are defined and their algebraic and topological properties discussed. The relationship between one of these classes of matrices and an important class of linear transformations on a sequence space is derived and discussed. The treatment here leans heavily on that of Cooke [1][1] for theorems on infinite matrices. Only those results of use in the applications of the subsequent chapters are set down; as a consequence, the analysis is concise and of an elementary nature.

Definitions.--An infinite matrix is an array $A = (a_{ij})$, $i,j=1,2,3, \ldots$, of complex numbers with matrix addition and matrix and scalar multiplication defined by

$$A+B = (a_{ij} + b_{ij}) \tag{1.1}$$

$$AB = (\sum_{k=1}^{\infty} a_{ik}b_{kj}) \tag{1.2}$$

$$\lambda A = (\lambda a_{ij}) \tag{1.3}$$

where, in (1.2), $\sum_{k=1}^{\infty} a_{ik}b_{kj}$ is assumed to exist for every i,j.

We shall need two main classes of infinite matrices: the lower semi-matrices (l.s.m.) and the diagonally invariant matrices (d.i.m.). A matrix $A = (a_{ij})$ is said to be a l.s.m. if and only if

---

[1] Numbers in square brackets refer to the List of References preceding the Appendix.

$$a_{ij} = 0 \text{ for } j > i$$

while A is a d.i.m. if and only if it is a l.s.m. and

$$a_{ij} = \alpha_{i-j} \text{ for every } i, j.$$

Thus, pictorially, we have for a l.s.m. A

$$A = \begin{pmatrix}
a_{11} & 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots \\
a_{21} & a_{22} & 0 & 0 & \cdots & \cdots & 0 & \cdots \\
a_{31} & a_{32} & a_{33} & 0 & \cdots & \cdots & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n,n-1} & a_{nn} & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}$$

and, if A is a d.i.m,

$$A = \begin{pmatrix}
\alpha_0 & 0 & 0 & 0 & \cdots & \cdots & 0 & \cdots \\
\alpha_1 & \alpha_0 & 0 & 0 & \cdots & \cdots & 0 & \cdots \\
\alpha_2 & \alpha_1 & \alpha_0 & 0 & \cdots & \cdots & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
\alpha_{n-1} & \alpha_{n-2} & \alpha_{n-3} & \cdots & \alpha_1 & \alpha_0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}$$

We shall denote the set of all l.s.m. by $\mathcal{L}$ and the set of all d.i.m. by $\mathcal{D}$.

Denoting by σ the linear vector space of all sequences with complex entries, the equations

$$y_n = \sum_{k=0}^{n} a_{nk}x_k, \; n = 0,1,2, \ldots, \qquad (1.4)$$

where a_{nk}, y_n, x_n are complex numbers, constitute a linear transformation of σ on itself. When $a_{nk} = b_{n-k}$, the transformation determined by (1.4) will be called a constant linear filter; otherwise, a variable linear filter.[2]

Using any one of the following notations

$$u = (u_0, u_1, u_2, \ldots) = (u_n) \qquad (1.5)$$

for (row) vectors in σ, (1.4) may be written as the vector-matrix equation

$$y^T = Ax^T \qquad (1.6)$$

where $A \varepsilon \mathcal{L}$ when (1.4) represents a variable filter and $A \varepsilon \mathcal{D}$ when (1.4) represents a constant filter. The superscript T denotes the transposed vector. Hereafter, this notation will be dropped since it will be obvious from the context whether a row or column vector is intended.

Algebraic character of $\mathcal{L}$ and $\mathcal{D}$. Reciprocals.--$\mathcal{L}$ and $\mathcal{D}$ are completely characterized algebraically by the following statement: with the operations defined in (1.1), (1.2), (1.3), $\mathcal{L}$ is a non-Abelian algebra with unit element over the complex field and $\mathcal{D}$ is an Abelian sub-algebra, with unit element, of $\mathcal{L}$. Moreover, $\mathcal{D}$ contains no proper divisors of zero. The definition of "algebra" used here is that of [3], pp. 19, 20. For the convenience of the reader, the definition is reproduced in the Appendix. Cooke [1], p. 7 ff., in effect has shown that $\mathcal{L}$ is a non-Abelian ring with unit element while Matthews [2] has shown that $\mathcal{D}$ is an Abelian ring with unit element but without proper divisors of zero (integral domain). The remaining conditions that $\mathcal{L}$ and $\mathcal{D}$ must satisfy in order to be classed as algebras are immediate consequences of the definitions (1.1), (1.2) and (1.3).

---

[2] Calling such transformations filters anticipates the application to be made in the next chapter, for control engineers frequently refer to a linear transmission system as a filter [7].

On the basis of theorems proved for l.s.m. in [1], ch. 2, we may state that, if $A \varepsilon \mathcal{L}$ and $a_{ii} \neq 0$ for every i, then A has a unique (two-sided) reciprocal, A^{-1}, in $\mathcal{L}$ whose principal diagonal elements are equal to $1/a_{ii}$.[3] Moreover, if $a_{ii} = 0$ for some i, then $A \varepsilon \mathcal{L}$ has no reciprocal in $\mathcal{L}$, either left-hand or right-hand. It follows from [1], p. 25, that, if $A, B \varepsilon \mathcal{L}$ are non-singular, then $B^{-1}A^{-1} = (AB)^{-1}$ and $A^{-1}(I + BA^{-1})^{-1} = (A + B)^{-1}$.

A useful algorithm for the determination of the reciprocal of $A \varepsilon \mathcal{L}$ is readily established. Let $B = (b_{ij}) \varepsilon \mathcal{L}$ be the unique reciprocal of A. Then the k-th row of the product BA equals the k-th row of the unit matrix. Clearly, the nonzero components of the k-th row of BA can be obtained from the finite vector-matrix equation

---

[3] B is said to be a two-sided reciprocal of A if both $AB = I$ and $BA = I$.

$$(b_{k1}\ b_{k2}\ \cdots\ b_{kk}) \begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kk} \end{pmatrix} = (0\ 0\ .\ .\ .\ 0\ 1). \tag{1.7}$$

Denoting the finite matrix of (1.7) by A_k, we have the well-known result

$$A_k^{-1} = \operatorname{adj} A_k / \det A_k \tag{1.8}$$

where adj A_k and det A_k, as defined in [4], ch. 4, are respectively the matrix adjoint to A_k and the determinant of A_k.

Evidently

$$\det A_k = \prod_{l=1}^{k} a_{ll}. \tag{1.9}$$

Letting adj $A_k = (l_{ij}^k)$, we clearly need only determine l_{kj}^k, $j=1,2,\ldots,k$. But

$$\begin{aligned} l_{kj}^k &= \text{cofactor } a_{jk} \\ &= (-1)^{k-j} \prod_{l=1}^{j-1} a_{ll} \det \omega_{kj} \end{aligned}$$

where

$\det \omega_{kk} = 1$

$$\det \omega_{kj} = \begin{vmatrix} a_{j+1,j} & a_{j+1,j+1} & 0 & 0 & \cdots & 0 & 0 \\ a_{j+2,j} & a_{j+2,j+1} & a_{j+2,j+2} & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{k-1,j} & a_{k-1,j+1} & a_{k-1,j+2} & & \cdots & a_{k-1,k-2} & a_{k-1,k-1} \\ a_{kj} & a_{k,j+1} & a_{k,j+2} & & \cdots & a_{k,k-2} & a_{k,k-1} \end{vmatrix}, j=1,2,\ldots,k-1.$$

Thus

$$b_{kj} = (-1)^{k-j} \det \omega_{kj} / \prod_{l=j}^{k} a_{ll} \quad \begin{aligned} j &= 1,2,\ldots k \\ k &= 1,2,3,\ldots\ . \end{aligned} \tag{1.10}$$

Alternatively, a recursion scheme results directly and simply from (1.7):

$$b_{kj} = \frac{-1}{a_{jj}}[a_{kj}b_{kk}+a_{k-1,j}b_{k,k-1}+\dots+a_{j+1,j}b_{k,j+1}], \quad 1 \leq j \leq k-1. \qquad (1.11)$$

If $A \epsilon \mathcal{D}$, then $a_{ij} = \alpha_{i-j}$ and (1.10) becomes

$$b_{kj} = (-1)^{k-j} \det \mu_{k-j}/\alpha_0^{k-j+1} \qquad (1.12)$$

where

$$\det \mu_0 = 1$$

$$\det \mu_m = \begin{vmatrix} \alpha_1 & \alpha_0 & 0 & 0 & \dots & 0 \\ \alpha_2 & \alpha_1 & \alpha_0 & 0 & \dots & 0 \\ . & . & . & . & . & . \\ \alpha_m & \alpha_{m-1} & \alpha_{m-2} & \dots & \alpha_2 & \alpha_1 \end{vmatrix}, \quad m = 1,2,\dots,k-1. \qquad (1.13)$$

Let us now introduce the d.i.m. $D_n = (d_{ij}^n)$ defined by

$$d_{ij}^n = \delta_{i-j+1,n+1}, \quad n = 0,1,2,\dots, \qquad (1.14)$$

where δ_{jk} is the Kronecker delta. Pictorially we have

$$D_n = \begin{pmatrix} 0_n \\ I \end{pmatrix}$$

where the notation on the right indicates a row partition of the matrix in which the upper submatrix consists of n rows of zeros and the lower submatrix is the unit matrix. Thus $D_0 = I$, the unit matrix, and it is easy to see that D_n satisfies

$$D_m \cdot D_n = D_{m+n}, \quad m,n = 0,1,2,\dots \qquad (1.15)$$

and, hence, that

$$(D_m)^n = D_{mn}, \quad m,n = 0,1,2,\dots .$$

If $C \in \mathcal{D}$ and $C \neq 0$, then it is evident that there exists a unique non-negative integer n such that

$$C = D_n A$$

where A is unique and non-singular. If $n \neq 0$, then A will be called the pseudo-reciprocal of C and it follows that

$$CA^{-1} = A^{-1}C = D_n.$$

Pseudo-reciprocals can be defined in the same fashion for some matrices in the complement of $\mathcal{D}$ in $\mathcal{L}$.

A useful property of the D_n appears in the following

Theorem I. Let $CX = E$ where $C = D_n A$, $E = D_m B$ with $A, B \in \mathcal{D}$ non-singular. Then a necessary and sufficient condition (n.a.s.c.) that the equation have a unique solution X in $\mathcal{D}$ is that $m \geq n$.

The condition is sufficient, for

$$D_n AX = D_m B$$

and, multiplying on the left by A^{-1} and applying (1.15) and the commutative property,

$$D_n X = D_{m-n} D_n A^{-1} B.$$

Finally, $$X = D_{m-n} A^{-1} B,$$

since the cancellation laws hold in $\mathcal{D}$.

For the necessity, let $X = D_k F$, where F is non-singular. Then from the given equation

$$D_n D_k AF = D_m B$$

$$D_{n+k} = D_m B(AF)^{-1}.$$

Hence, we must have $B(AF)^{-1} = D_p$. But $B(AF)^{-1}$ is itself non-singular; consequently $p = 0$. From the resulting equation,

$$D_{n+k} = D_m,$$

the theorem follows.

Convergence and Summability in σ, $\mathcal{L}$ and $\mathcal{B}$.--Let

$$x^k = (x_0^k, x_1^k, \ldots, x_n^k, \ldots),\ k = 0,1,2,\ldots,$$

be a sequence of vectors in σ. The sequence will be said to converge to the vector $x = (x_0, x_1, x_2, \ldots, x_n, \ldots)$--written

$$\sigma\text{-}\lim_{k\to\infty} x^k = x$$

--if and only if

$$\lim_{k\to\infty} x_n^k = x_n,\ n = 0,1,2,\ldots .$$

Let $A^k = (a^{k0}, a^{k1}, \ldots, a^{kn}, \ldots)$, $k = 1,2,3,\ldots$, be a sequence of matrices in $\mathcal{L}$, where a^{kj} denotes a column vector--in σ-- of A^k. Then the sequence A^k, $k = 1,2,3,\ldots$, will be said to converge to the matrix $A = (a^0, a^1, \ldots, a^n, \ldots) \varepsilon \mathcal{L}$--written

$$\mathcal{L}\text{-}\lim_{k\to\infty} A^k = A$$

--if and only if

$$\sigma\text{-}\lim_{k\to\infty} a^{kn} = a^n,\ n = 0,1,2,\ldots .$$

An analogous definition obtains for

$$\mathcal{B}\text{-}\lim_{k\to\infty} A^k.$$

These limit processes are examples of what Cooke [1], p. 283, calls coordinate (c-) convergence. It can be shown, although we shall not need the fact here, that metrics can be defined in σ and $\mathcal{L}$ in such a way that the resulting metric spaces are topologically equivalent to the spaces σ and $\mathcal{L}$ with the just-defined topology.[4]

It is apparent that the mapping which assigns to each $A = (\alpha_{i-j}) \varepsilon \mathcal{D}$ the vector $\alpha = (\alpha_0, \alpha_1, \ldots) \varepsilon \sigma$ is a homeomorphism. The correspondence is preserved moreover by all the algebraic operations for $\mathcal{D}$ and σ, where, in the latter, multiplication is defined by

$$(\alpha_n)(\beta_n) = (\sum_{k=0}^{n} \alpha_{n-k}\beta_k).$$

Let

$$y = Ax$$

where A is an arbitrary infinite matrix, not necessarily in $\mathcal{L}$. A will be called a Kojima (K-) matrix whenever the following conditions are satisfied [1]:

$$\text{i)} \quad \sum_{k=1}^{\infty} |a_{nk}| \leq M \text{ for every } n$$

$$\text{ii)} \quad \lim_{n \to \infty} a_{nk} = \alpha_k \text{ for every fixed } k \qquad (1.16)$$

$$\text{iii)} \quad \lim_{n \to \infty} \sum_{k=1}^{\infty} a_{nk} = \alpha.$$

The numbers α, α_k are called the characteristic numbers of the K-matrix. If $\alpha_k = 0$, we shall call A a T_α-matrix; a T_1-matrix is usually called a Toeplitz (T-) matrix. If (1.16,i) alone is satisfied, A is called a K_r-matrix.

---

[4] For an illustration of this fact for σ, see Cooke [1], p. 315. The demonstration for $\mathcal{L}$ is a commonplace extension of the concept of metric.

Two theorems on K and K_r matrices are important in the present connection. These appear in [1], pp. 63 and 299, and are presented here without proof.

Theorem II (Kojima-Schur). A n.a.s.c. that the sequence

$$y_n = \sum_{k=1}^{\infty} a_{nk}x_k, \quad n = 1,2,3,\dots , \tag{1.17}$$

converge when x_n does is that (a_{ij}) be a K-matrix. Moreover $\lim_{n\to\infty} x_n = \bar{x}$ implies that

$$\bar{y} = \lim_{n\to\infty} y_n = \alpha\bar{x} + \sum_{k=1}^{\infty} \alpha_k(x_k-\bar{x}). \tag{1.18}$$

Theorem III (Köthe-Toeplitz). Suppose

$$\sum_{k=1}^{\infty} |a_{nk}| < \infty \tag{1.19}$$

for every n. Then a n.a.s.c. that y_n of (1.17) be bounded when x_n is, is that (a_{ij}) be a K_r-matrix.

Clearly, when $A \varepsilon \mathcal{L}$ (1.19) is satisfied. Thus, denoting by σ_∞ the space of all bounded sequences and calling a l.s.m. stable when it maps σ_∞ on itself, we may state

Theorem IV. $A \varepsilon \mathcal{L}$ is stable if and only if it is a K_r-matrix.

Corollary IV. Let $A = (\alpha_{i-j}) \varepsilon \mathcal{D}$. Then A is stable if and only if

$$\sum_{k=0}^{\infty} |\alpha_k| < \infty .$$

Apparently, a stable d.i.m. is a T_α-matrix.

Displacement transformations on σ.--The set of all matrix polynomials of the form

$$F(D) = a_0 I + a_1 D_1 + \dots + a_n D_n \tag{1.20}$$

$$= a_0 I + a_1 D_1 + \dots + a_n (D_1)^n$$

forms a commutative ring with unit element. The proof is similar to that given in [4], p. 157 ff. It is a trivial matter to show that the set of all such polynomials is dense in $\mathscr{D}$, in the topology established for $\mathscr{D}$ in a previous section.

Using the equation

$$D_1(x_0, x_1, x_2, \dots) = (0, x_0, x_1, \dots) \tag{1.21}$$

as a guide, let us define a transformation, E^{-1}, on σ as

$$E^{-1}(x_0, x_1, x_2, \dots) = (0, x_0, x_1, \dots). \tag{1.22}$$

From (1.22) we have, recursively,

$$(E^{-1})^k\, x = E^{-1}(E^{-1})^{k-1}\, x, \ k = 2, 3, 4, \dots\ . \tag{1.23}$$

If we define $(E^{-1})^0 = 1$, then (1.23) holds for $k = 1$ and, to simplify the notation, we define

$$E^{-k} = (E^{-1})^k, \ k = 1, 2, 3, \dots\ . \tag{1.24}$$

As with the matrix polynomials of (1.20), it can be shown that the set of all polynomials

$$a_0 + a_1 E^{-1} + \dots + a_n E^{-n} \tag{1.25}$$

over the complex field forms a commutative ring with unit element.

Now let (v_n) be a fixed vector and let a_r, $r = 0, 1, 2, \dots$, be an arbitrary sequence of fixed complex numbers. Then the sequence of vectors u^m defined by

$$u^m = \{\sum_{r=0}^{m} a_r E^{-r}\}(v_n), \ m = 0,1,2,\ldots \tag{1.26}$$

is such that

$$\sigma\text{-}\lim_{m\to\infty} u^m = u$$

where

$$u_n = u_n^m \quad \text{for } n \leq m.$$

Equations (1.26) thus uniquely determine a linear transformation of (v_n) onto (u_n). We may say that the sequence of transformations

$$\sum_{r=0}^{m} a_r E^{-r}, \ m = 0,1,2,\ldots, \tag{1.27}$$

converges to a transformation which we shall denote, with an obvious interpretation, by

$$\sum_{r=0}^{\infty} a_r E^{-r}. \tag{1.28}$$

We shall call such transformations displacement transformations.

It follows from our construction that the product of two such transformations is just the Cauchy product

$$\{\sum_{r=0}^{\infty} a_r E^{-r}\}\{\sum_{r=0}^{\infty} b_r E^{-r}\} = \sum_{r=0}^{\infty}\{\sum_{j=0}^{r} a_j b_{r-j}\}E^{-r}.$$

Also, under the topology introduced it is easy to see that, as is the case in $\mathscr{D}$, the set of all polynomials (1.25) is dense in the space of all displacement transformations. The correspondence

$$D_n \leftrightarrow E^{-n} \tag{1.29}$$

then establishes a bicontinuous isomorphism of $\mathscr{D}$ onto the space of all displacement transformations.

The notation $(u_0,u_1,u_2,\ldots)$ in (1.5) for vectors in σ presupposes a particular choice of a basis for σ relative to which $u_0,u_1,u_2,\ldots$ are the coordinates of u. The set of vectors

$$\varepsilon^r = (\delta_{nr}),\ r = 0,1,2,\ldots, \tag{1.30}$$

where δ_{nr} is the Kronecker delta forms such a basis, in terms of which (u_n) has the unique representation

$$(u_n) = \sum_{r=0}^{\infty} u_r\varepsilon^r. \tag{1.31}$$

Applying the previously deduced convergence properties of displacement transformations to (1.31) we obtain

$$(u_n) = \{\sum_{r=0}^{\infty} u_rE^{-r}\}\varepsilon^0 . \tag{1.32}$$

Thus, if we adhere to use of the ε-basis--as we shall in the sequel--we see that the study of constant linear filters on σ of the form

$$y_n = \sum_{k=0}^{n} a_{n-k}x_k \tag{1.33}$$

is equivalent to study of displacement transformation equations of the form

$$\{\sum_{r=0}^{\infty} y_rE^{-r}\}\varepsilon^0 = \{\sum_{r=0}^{\infty} a_rE^{-r}\}\{\sum_{r=0}^{\infty} x_rE^{-r}\}\varepsilon^0 \tag{1.34}$$

or of the corresponding matrix equations in $\mathscr{B}$. The retention of the vector ε^0 in (1.34) is evidently superfluous.

Under the isomorphism established by (1.29), multiplicative inverses correspond so that the coefficients of the serial expansion of the inverse of (1.28)--assuming its existence--are given by (1.12).

The same result is obtained [5] by carrying out the formal reciprocation

$$[\sum_{r=0}^{\infty} a_r E^{-r}]^{-1} = 1/\sum_{r=0}^{\infty} a_r E^{-r}. \tag{1.35}$$

We could now carry out a discussion of pseudo-inverses of singular displacement transformations, but, since the analogy with pseudo-reciprocals is precise, the discussion will be omitted.

By virtue of the foregoing considerations, every Maclaurin's series

$$\sum_{r=0}^{\infty} a_r z^{-r} \tag{1.36}$$

where z^{-1} is a complex variable becomes a valid displacement transformation when z^{-1} is replaced by E^{-1}. This is true even if (1.36) converges only trivially.

If (1.36) should happen to have a non-trivial circle of convergence and, as a consequence, uniquely represent an analytic function within that circle, and if, in addition, the function so represented can be expressed in closed form, then a new definition, convenient for the applications, suggests itself.

Let $A(z^{-1})$ denote the closed form--now assumed to exist--of the function (1.36).[5] Then we define the transformation $A(E^{-1})$ as

$$A(E^{-1}) = \sum_{r=0}^{\infty} a_r E^{-r}. \tag{1.37}$$

---

[5] Or, more generally, we can let $A(z^{-1})$ represent the analytic configuration obtained by continuing (1.36) along every path. (Cf.,[6]).

On the other hand, any function $A(z^{-1})$ analytic in a neighborhood of $z^{-1} = 0$ leads to a transformation $A(E^{-1})$ which will be defined to be equal to its Maclaurin's expansion. In this case $A(E^{-1})$ is well defined since, by Cauchy's integral formula, the coefficients a_r are uniquely determined by

$$a_r = \frac{1}{r!} A^{(r)}(0) = \frac{1}{2\pi i}\int_C \frac{A(\zeta)}{\zeta^{r+1}}\, d\zeta \qquad (1.38)$$

where C is a simple closed curve about $z^{-1} = 0$ which excludes the singularities of $A(z^{-1})$. Even more important for the applications is the fact that the series in (1.37), when $A(E^{-1})$ is a fractional rational transformation, can be obtained by division of the numerator by the denominator of $A(E^{-1})$.

As will be seen, the most important class of fractional rational displacement transformations--from the standpoint of applications at least--is that given by

$$\Psi_n(aE^{-1}) = \sum_{r=1}^{\infty} r^n a^r E^{-r} \qquad (1.39)$$

$$= \frac{S_n(aE^{-1})}{(1-aE^{-1})^{n+1}}, \quad n = 0,1,2,\ldots, \qquad (1.40)$$

where a is complex and

$$S_0(z) = z$$

$$S_n(z) = \sum_{q=1}^{n} \sum_{m=0}^{q-1} (-1)^m (n+1;m)(q-m)^n z^q, \quad n = 1,2,3,\ldots, \qquad (1.41)$$

with $(n;j) = n!/[j!(n-j)!]$. The properties of the functions $\Psi_n(z)$ and $S_n(z)$ have been investigated in detail by Lawden [7].

The stability criterion of Corollary IV has a convenient expression for displacement transformations. Let $\sum_{r=0}^{\infty} \alpha_r E^{-r}$ be the image of the d.i.m. A of the corollary under the mapping (1.29). Then it can be shown by a proof precisely similar to that of Hurewicz [6] that A is stable if and only if $\sum_{r=0}^{\infty} a_r z^{-r}$ has no singularities within or on the circle $|z^{-1}| = 1$; A will be unstable if there is at least one singularity in $|z^{-1}| < 1$. The case of a singularity on $|z^{-1}| = 1$ represents a condition of marginal stability.

CHAPTER II

FORMULATION OF THE THEORY OF SAMPLED-DATA CONTROL

Introduction.--As reference to the pertinent literature--much of which is cataloged in [8], [9] and [10]--reveals, a relatively rich theory of sampled-data control systems has been developed since the initial work of MacColl [11] and Hurewicz [6]. Extensive use of what has come to be known as the "z-transform" has imparted to this theory an operational character analogous to that of the theory of continuous-data control systems wherein Laplace transforms are utilized.

While the utility and power of z-transform methods are well substantiated, the fundamental formulation of the z-transform concept is open to question by the mathematician. Indeed, control engineers themselves are by no means in complete agreement on this point. Thus, while the complex series

$$\sum_{r=0}^{\infty} f(rh)z^{-r}$$

--where $f(rh)$, $r = 0,1,2,\ldots$, are the successive values of a real-valued function $f(t)$ of the real variable t and h is a small positive constant--is usually referred to as the z-transform, it is not agreed in existing treatments whether it is to be regarded as the transform of the function $f(t)$ or only of the sequence $\{f(rh)\}$. The question is important, for in the former case a unique inverse transform does not exist while, in the latter, a unique inverse does exist.

In some instances, the series appears as the result of Laplace transformation of the series

$$\sum_{r=0}^{\infty} f(t)\delta(t-rh)$$

and a subsequent change of variable. Here, $\delta(t)$ denotes Dirac's impulse function. The questions of rigor accruing to the use of the Dirac function in the above connection have not, to my knowledge, been answered in the literature of sampled-data control.

A consideration of the applications of the z-transform makes evident the fact that its principal virtue is the facility it confers upon the carrying out of certain algebraic operations with sequences of function values. Moreover, the fact that z^{-1} is a symbol for a complex variable is of no particular importance for the applications except perhaps in considerations of stability. This latter advantage seems hardly to compensate for the questions of convergence raised by retention of the complex function form.

These facts, together with certain indications in the work of Hurewicz [6], Friedland [12] and Zadeh [13], suggest a treatment of sampled-data systems as linear transformations on a sequence space. In this chapter such a treatment is developed utilizing the theory expounded in the first chapter.

The present approach provides a unified treatment of both time-invariant and time-varying sampled-data systems, free of the Dirac function and of superfluous dependence on complex function theory;

both the "classical" and Jury's "modified" z-transform [14] appear unimpaired for engineering applications, with questions relating to inverse transforms resolved.

Sampled-data systems as linear filters on σ.--Sampled-data systems generally occur in two distinct ways in practice. One is exemplified by the digital computer which operates--linearly or nonlinearly--on a sequence of numbers, the input, to produce another sequence of numbers, the output. The other is exemplified by the continuous-data system, such as a servomechanism, operating on a piecewise constant input, where the discontinuities of the input are usually assumed to be equidistantly spaced. We shall consider here the application of the matrix theory developed in the preceding sections only to the second type of (linear) sampled-data system alluded to above.

The set $\mathcal{F}$ of all piecewise continuous complex valued functions defined on the non-negative real axis and zero elsewhere, forms a linear vector space over the complex field, with obvious definitions of addition and multiplication by a scalar. Now let f^* denote the set of all vectors $(f_k^\tau)\varepsilon\sigma$ for which

i) f_k^τ is a piecewise continuous complex valued function of τ for $\tau\varepsilon[0,1]$, $k = 0,1,2,\dots$;

ii) $f_{k+1}^0 = f_k^1$, $k = 0,1,2,\dots$.

Then the class $\mathcal{F}^*$ of all sets f^* is a linear vector space over the complex field with addition and scalar multiplication defined by

$$f^* + g^* = \{(p_n^\tau) \mid p_n^\tau = f_n^\tau + g_n^\tau\};$$
$$cf^* = \{(p_n^\tau) \mid p_n^\tau = cf_n^\tau\}.$$

Now if, for $f^* \varepsilon\, \mathcal{F}^*$, we put

$$f_k^\tau = \begin{cases} f[(k+\tau)h], & 0 \le \tau \le 1 \\ 0 & \text{, elsewhere} \end{cases} \tag{2.1}$$

where $f(t) \varepsilon\, \mathcal{F}$ and h is a positive constant, then the mapping

$$f(t) \longrightarrow f^* \tag{2.2}$$

is an isomorphism of $\mathcal{F}$ onto $\mathcal{F}^*$ while the mapping

$$f(t) \longrightarrow (f_n^\tau) \tag{2.3}$$

for a fixed value of $\tau \varepsilon [0,1]$ is a homomorphism of $\mathcal{F}$ onto σ.

A linear automatic control system with input $f(t)$ and output $y(t)$ is described, under wide conditions, by an integral equation of the form [13]

$$y(t) = \int_0^t w(t,\mu) f(\mu)\, d\mu \tag{2.4}$$

where $w(t,\tau)$--frequently referred to as Green's function--is a continuous function of $\tau \varepsilon [0,t]$ for every positive t and $w(t,\tau) = 0$ for $\tau > t$. When the parameters of the system are constant in time, then $w(t,\tau)$ has the form

$$w(t,\mu) = v(t-\mu). \tag{2.5}$$

Thus (2.4) is a linear transformation of $\mathcal{F}$ into itself.

If we choose a value of $h > 0$ and define $f(t)$ such that

$$f(t) = f_k = \text{constant},\ kh \le t < (k+1)h,\ k = 0,1,2,\ldots, \tag{2.6}$$

then the set $\mathcal{J} \subset \mathcal{F}$ of such functions is mapped linearly into $\mathcal{F}$ by (2.4) and we have a sampled-data system of the second type discussed at the outset. Indeed, using (2.6) and putting $t = (n+\tau)h$, $0 \leq \tau \leq 1$, (2.4) can be written as

$$y[(n+\tau)h] = \sum_{k=0}^{n} f_k \int_{kh}^{(k+1)h} w[(n+\tau)h,\mu]d\mu \quad (2.7)$$

where account has been taken that

$$\int_{nh}^{(n+\tau)h} w[(n+\tau)h,\mu]d\mu = \int_{nh}^{(n+1)h} w[(n+\tau)h,\mu]d\mu.$$

Hence, putting

$$c_{nk}^{\tau} = \int_{kh}^{(k+1)h} w[(n+\tau)h,\mu]d\mu, \quad (2.8)$$

(2.7) becomes

$$y[(n+\tau)h] = \sum_{k=0}^{n} c_{nk}^{\tau} f_k. \quad (2.9)$$

Using (2.1) and (2.2) we obtain y^* as the set of all vectors in σ of the form

$$(y_n^{\tau}) = (\sum_{k=0}^{n} c_{nk}^{\tau} f_k), \quad (2.10)$$

a variable linear filter.

For the case (2.5), we have

$$(y_n^{\tau}) = (\sum_{k=0}^{n} \gamma_{n-k}^{\tau} f_k), \quad (2.11)$$

a constant linear filter on σ, where

$$\gamma_n^{\tau} = \int_{(n-1+\tau)h}^{(n+\tau)h} v(\lambda)d\lambda. \quad (2.12)$$

Thus it is seen that the analysis of linear sampled-data systems is coextensive with the analysis of linear filters on σ. In the sequel, the matrices $(c^{\tau}_{i-1,j-1})$ and (γ^{τ}_{i-j}) will be denoted respectively by $C(\tau)$ and $\Gamma(\tau)$, while the displacement transformation corresponding to $\Gamma(\tau)$ under the mapping (1.29) will be denoted by $\Gamma(\tau;E^{-1})$. Hence (2.10) and (2.11) may be written, respectively, as

$$(y^{\tau}_{n}) = C(\tau)(f_{n}) \tag{2.13}$$

$$(y^{\tau}_{n}) = \Gamma(\tau)(f_{n}). \tag{2.14}$$

By virtue of (1.29), (2.14) has the equivalent expression

$$Y(\tau;E^{-1})\varepsilon^{0} = \Gamma(\tau;E^{-1})F(E^{-1})\varepsilon^{0} \tag{2.15}$$

where $Y(\tau;E^{-1})$ and $F(E^{-1})$ denote the displacement transformation representations (1.32) of (y^{τ}_{n}) and (f_{n}). It is now evident that $Y(0;E^{-1})$ is just the "classical" z-transform and $Y(\tau;E^{-1})$ Jury's "modified" z-transform [14].

The first type of sampled-data system, i.e., that exemplified by the digital computer, is described by a difference equation of the form [15]

$$a_{0}u_{n}+a_{1}u_{n-1}+\dots+a_{m}u_{n-m} = b_{0}u_{n}+b_{1}v_{n-1}+b_{2}v_{n-2}+\dots+b_{s}v_{n-s} \tag{2.16}$$

where the sequences $\{u_{n}\}$ and $\{v_{n}\}$ are respectively the output and input of the computer and the a_{i}, b_{i} are constants. Since the analysis of such systems by the present methods involves no concepts not previously discussed, a treatment of them will be dispensed with.

Explicit form of constant linear filters.--As is well known [16] the Green's function $v(t)$ of (2.5) consists of a linear sum, with complex coefficients, of terms of the form

$$t^{m}e^{\alpha t} \tag{2.17}$$

where m is a non-negative integer and α is a complex constant. Consequently, the coefficients γ^{τ}_{n} (2.12) of the transformation $\Gamma(\tau;E^{-1})$

of (2.15) are linear sums of terms of the form

$$(n+\tau)^m(e^{\alpha h})^n \tag{2.18}$$

so that it may be concluded that $\Gamma(\tau;E^{-1})$ is a linear sum of displacement transformations

$$\sum_{r=0}^{\infty} (r+\tau)^m(e^{\alpha h})^r E^{-r}. \tag{2.19}$$

The functions

$$\Phi_n(\tau;z) = \sum_{r=0}^{\infty} (r+\tau)^n z^r, \quad n = 0,1,2,\ldots, \tag{2.20}$$

are rational functions of z, as the following system of equations shows:

$$\Phi_n(\tau;z) = \tau^n + \sum_{r=1}^{\infty} \{ \sum_{j=0}^{\infty} (n;j)\tau^j r^{n-j} \} z^r \tag{2.21}$$

$$= \tau^n + \sum_{j=0}^{n} (n;j)\tau^j \{ \sum_{r=1}^{\infty} r^{n-j} z^r \} \tag{2.22}$$

$$= \tau^n + \sum_{j=0}^{n} (n;j)\tau^j \Psi_{n-j}(z) \tag{2.23}$$

$$= \{ \tau^n(1-z)^{n+1} + \sum_{j=0}^{n} (n;j)\tau^j(1-z)^j S_{n-j}(z) \} / (1-z)^{n+1} \tag{2.24}$$

$$= R_n(\tau;z)/(1-z)^{n+1}. \tag{2.25}$$

Inspection of (2.24) reveals that $R_n(\tau;z)$ is a polynomial of degree n in z, with $R_n(0;z) = S_n(z)$. An explicit form for the coefficients of the powers of $R_n(\tau;z)$ may be obtained by using (1.41) in (2.24). However, it is more direct to write

$$R_n(\tau;z) = \sum_{m=0}^{n} b_m^n(\tau)z^m = (1-z)^{n+1} \sum_{r=0}^{\infty} (r+\tau)^n z^r. \qquad (2.26)$$

There results

$$b_0^n(\tau) = \tau^n$$

$$b_m^n(\tau) = (-1)^m(n+1;m)\tau^n + \sum_{r=0}^{n} (n;r)\tau^r \sum_{q=0}^{m-1} (-1)^q(n+1;q)(m-q)^{n-r}, \qquad (2.27)$$

$$m = 1,2,\ldots,n.$$

From the foregoing considerations, we deduce that $\Gamma(\tau;E^{-1})$ is a rational function of E^{-1} for every $\tau\varepsilon[0,1]$. On the other hand, it is a commonplace [16] that $v(t)$ has a rational Laplace transform. By virtue of (2.2) these facts establish an isomorphism between rational Laplace transforms and rational displacement transformations $\Gamma(\tau;E^{-1})$. For a fixed value of τ, the correspondence is merely a homomorphism, by virtue of (2.3).

(2.19) and (2.25) imply that the stability character of the constant linear filter represented by $\Gamma(\tau;E^{-1})$ is determined for any value of $\tau\varepsilon[0,1]$ by that character for a fixed value of τ in the interval; i.e., the poles of $\Gamma(\tau;E^{-1})$ are independent of τ. The criterion of stability for continuous-data control systems [16] requires that, for stability, the real parts of all the parameters α (2.17) appearing in $v(t)$ must be negative; one or more positive real parts imply instability and one or more zero real parts imply marginal stability. From these facts, together with (2.19) and (2.25), we may infer that, according as the continuous-data system is stable, unstable or marginally stable, its sampled-data counterpart possesses the same property and conversely.

Error-sampled feedback systems.--In the simplest instance of a system of this type, (2.13) becomes

$$(y_n^\tau) = C(\tau)(f_n - y_n^0). \qquad (2.28)$$

For $\tau = 0$, we have from (2.28),

$$(y_n^0) = [I + C(0)]^{-1}C(0)(f_n). \qquad (2.29)$$

The inverse of $I + C(0)$ always exists, since by (2.08)

$$c_{nn}^0 = 0, \ n = 0,1,2,\dots \ .$$

Substituting (2.29) in (2.28) yields

$$\begin{aligned}(y_n^\tau) &= C(\tau)\{I - [I + C(0)]^{-1}C(0)\}(f_n) \qquad (2.30)\\ &= C(\tau)[I + C(0)]^{-1}(f_n).\end{aligned}$$

That (2.30) reduces to (2.29) for $\tau = 0$ follows from the identity, readily proven for $\mathcal{L}$,

$$A(I + A)^{-1} = (I + A)^{-1}A. \qquad (2.31)$$

The case of error-sampling for constant linear filters leads to the revised form of (2.15)

$$Y(\tau;E^{-1})\varepsilon^0 = P(\tau;E^{-1})F(E^{-1})\varepsilon^0 \qquad (2.32)$$

where

$$P(\tau;E^{-1}) = \frac{\Gamma(\tau;E^{-1})}{1 + \Gamma(0;E^{-1})} \ . \qquad (2.33)$$

On the basis of the discussion of the previous section, we may conclude that the only singularities of $P(\tau;E^{-1})$ are the roots of

$$\Gamma(0;z^{-1}) = -1. \qquad (2.34)$$

From the definition of stability for constant linear filters, it follows that, for $P(\tau;E^{-1})$ to represent a stable filter, the roots of (2.34) must lie outside the circle $|z^{-1}| = 1$ in the z^{-1}-plane.

The analysis of more complicated feedback systems differs only in detail from that given here.

Remarks.--We have indicated in the preceding sections a relatively extensive theory of constant linear filters based principally on an explicit general form for such filters. An analogous development cannot, unfortunately, be supplied for variable linear filters. Several factors contribute to this state of affairs: there is no explicit general form available for the l.s.m. $C(\tau)$ of (2.13). Moreover, since $\mathcal{D}$ is not dense on $\mathcal{L}$, $C(\tau)$ cannot be expressed as a limiting case of an infinite set of d.i.m. $C(\tau)$ may not have a pseudo-reciprocal; we have previously remarked on the noncommutativity of $\mathcal{L}$.

Friedland, who up to the present has published the only discussion--known to the present author--of variable linear filters [12], [17], has introduced a formalism for such filters by defining a "w-transform"; however, it is questionable whether the use of this transform offers any significant advantage in theory or practice over direct consideration of $C(\tau)$ itself.

Even the broad question of stability of such filters is left largely open. For, whereas the small investment of stability in a d.i.m. earns the profit of reproduction,[6] the same requirement for a l.s.m. may, as Friedland [17] has pointed out, result in a filter which, without further restriction, is not "stable" in any intuitively or practically acceptable sense. However it has not been the purpose here to investigate in depth the problem of variable linear filters per se, so the matter will be left at this point.

---

[6] I.e., the output of a stable constant linear filter is, ultimately, a reproduction of the input, up to a multiplicative constant (vide Theorem II).

Finally, we wish to call attention to the fact that the theory of sampled-data systems as developed here may be subsumed under Zadeh's general theory of linear filters [13]. Thus, where Zadeh presented a theory for "signal spaces" whose elements are images of a nondenumerable set, we have effectively extended the theory to cover signal spaces whose elements are images of a denumerable set.

CHAPTER III

NUMERICAL INTEGRATING OPERATORS

Introduction.--Tustin [18], [19], and others [20],...,[39] have in recent years developed a finite difference method for the approximate solution of ordinary differential equations whose application presupposes the availability of a certain set of linear (n-th order integrating) operators. These operators are characterized by the following property: let $f(t)$ be Riemann integrable on the interval $[0, t_0]$ and let h be a small positive constant. Then if P^{-n} denote an n-th order integrating operator, the result of the operation

$$P^{-n} \otimes [f(0), f(h), \ldots, f(mh)], \tag{3.1}$$

where $t_0 = mh$, is a sequence

$$F(0), F(h), \ldots, F(mh) \tag{3.2}$$

for which

$$F(kh) \approx \int_0^{kh} \int_0^{t^n} \cdots \int_0^{t^2} f(t^1)\,dt^1 \ldots dt^{n-1} dt^n, \quad k = 0,1,\ldots,m. \tag{3.3}$$

Several different sets of such operators have been introduced in the literature. One of these sets, first derived by Tustin, we shall refer to as Type I operators. Type II operators were derived by Ragazzini and Bergen, while a modified form of these, designated Type III, were introduced by Boxer and Thaler. Type IV operators were an innovation of Madwed while Cruickshank obtained Type V operators.

The aim of the cited references is utility, often at the expense of rigor. However, a general mathematical treatment, including an error analysis, of the operators of Types I, II and III has been given by Wasow [40]. Wasow's approach is via a certain discrete approximation to the Laplace integral.

In this chapter, operators of Types I, II and IV are developed as linear filters on σ_r, the space of all real sequences. This development, in addition to being of an inherently more elementary nature than Wasow's, clears up several questions of long standing with regard to application of the operators to integrands having nonzero initial values.

The application of numerical integrating operators to variable coefficient differential equations given in the sequel is new. The only previous application of the operators to such equations was by means of an algorithm suggested by Boxer and Thaler [33] and discussed in detail by Wasow. The present method, while no more difficult to apply than the Boxer-Thaler algorithm, is more accurate.

Tustin, as well as Boxer and Thaler, has attempted to apply numerical integrating operators to the approximate solution of non-linear differential equations. No such application is contemplated in the present chapter, for our methods are linear and it is well known from experience in many researches that the application of linear methods to non-linear problems is invariably considerably less than completely satisfactory.

Much of the material contained in this chapter has already been published as reference [41].

Elementary considerations.--In later sections we shall need several elementary results from the calculus and from interpolation theory. The first of these is the well-known identity

$$\int_0^x \int_0^{x^n} \cdots \int_0^{x^2} f(x^1)\,dx^1 \ldots dx^{n-1}\,dx^n = \int_0^x \frac{(x-s)^{n-1}}{(n-1)!} f(s)\,ds, \quad n=1,2,3,\ldots \,. \tag{3.4}$$

We shall consider three different approximations to the values of the integrals in (3.4), each leading to a different type of integrating operator.

Next, suppose that $f(t)$ is continuously differentiable on the non-negative real axis, that $f''(t)$ exists there and that the function and its derivatives are zero elsewhere. Then in each of the closed intervals $[t_{k-1}, t_k]$, $k = 1, 2, 3, \ldots$, with $t_0 = 0$, $f(t)$ has the representation [42]:

$$f(t) = \frac{f_k - f_{k-1}}{h}(t-t_{k-1}) + f_{k-1} + \frac{1}{2}(t-t_k)(t-t_{k-1})f''(\xi_k) \tag{3.5}$$

where $f(kh) = f_k$, $t_k - t_{k-1} = h =$ constant and $t_{k-1} < \xi_k < t_k$. It is observed that $\xi_k = \xi_k(t)$ and that (3.5) determines $f''(\xi_k(t))$ as a continuous function of t in the interior of the interval. Although $f''(\xi_k(t))$ is not defined at the endpoints of the interval, the right-hand (left-hand) limit of the function exists at the lower (upper) end point, so that with the expedient definition the function achieves continuity on the right (left). Thus if $g(t)$ is an integrable function which does not change sign on the interval $[a,b]$, where $t_{k-1} \leq a < b \leq t_k$, then by the first mean value theorem for integrals we have

$$\int_a^b f''(\xi_k(t))g(t)\,dt = f''(\eta_k)\int_a^b g(t)\,dt \tag{3.6}$$

where $t_{k-1} < \eta_k < t_k$.

Finally, we define the vector (p_m^n) as

$$(p_m^n) = (\int_0^{mh}\int_0^{t^n}\cdots\int_0^{t^2} f(t^1)dt^1\cdots dt^{n-1}dt^n),\ n = 1,2,3,\ldots, \tag{3.7}$$

and let

$$a_m^n = \int_{t_{m-1}}^{t_m}\int_0^{t^n}\cdots\int_0^{t^2} f(t^1)dt^1\cdots dt^{n-1}dt^n,\ m,\ n = 1,2,3,\ldots, \tag{3.8}$$

so that (3.7) may be written

$$(p_m^n) = (\sum_{k=1}^{m} a_k^n),\ n = 1,2,3,\ldots\ . \tag{3.9}$$

Alternately we may write, by virtue of (3.4),

$$(p_m^n) = (\int_0^{mh} \frac{(mh-s)^{n-1}}{(n-1)!} f(s)ds),\ n = 1,2,3,\ldots, \tag{3.10}$$

in which case we let

$$b_{m,k}^n = \int_{t_{k-1}}^{t_k} \frac{(mh-s)^{n-1}}{(n-1)!} f(s)ds,\ k,\ n = 1,2,3,\ldots\ . \tag{3.11}$$

Then it follows that (3.10) may be written

$$(p_m^n) = (\sum_{k=1}^{m} b_{m,k}^n)\ n = 1,2,3,\ldots\ . \tag{3.12}$$

Note that in (3.9) and (3.12), the summations appearing in the vectors on the right are to be set equal to zero for $m = 0$.

In the following analysis, we shall restrict our considerations to functions possessing a representation of the form (3.5).

Type I Operators.--From (3.8) and (3.9) we obtain

$$(p_m^1) = (\sum_{k=1}^{m} a_k^1) \tag{3.13}$$

and

$$a_k^1 = \int_{t_{k-1}}^{t_k} f(t)dt. \tag{3.14}$$

Substituting from (3.5) into (3.14) yields

$$a_k^1 = \int_{t_{k-1}}^{t_k} [\frac{1}{h}(f_k - f_{k-1})(t-t_{k-1}) + f_{k-1} + \frac{1}{2}(t-t_k)(t-t_{k-1})f''(\xi_k^1)]dt. \quad (3.15)$$

Noting that $\int_{t_{k-1}}^{t_k} \frac{1}{2}(t-t_k)(t-t_{k-1})dt = -\frac{h^3}{12}$ and using (3.6), we see that the last equation reduces to

$$a_k^1 = \frac{h}{2}(f_k + f_{k-1}) - \frac{h^3}{12} f''(\eta_k^1) \quad (3.16)$$

where $t_{k-1} < \eta_k^1 < t_k$.

Now substituting (3.16) into (3.13) we have

$$(p_m^1) = (\frac{h}{2} \sum_{k=0}^{m} (f_k + f_{k-1})) - \frac{h}{2}(f_0) - \frac{h^3}{12} (\sum_{k=0}^{m} f''(\eta_k^1)). \quad (3.17)$$

With (1.32) and the readily established relations

$$(\sum_{k=0}^{m} f_k) = \frac{1}{1-E^{-1}} (f_m) \quad (3.18)$$

$$(f_0) = \frac{f_0}{1-E^{-1}} \varepsilon^0 , \quad (3.19)$$

and the fact that $f_{-1} = 0$, (3.17) becomes

$$(p_m^1) = \frac{h}{2}\left[\frac{1+E^{-1}}{1-E^{-1}} F(E^{-1}) - \frac{f_0}{1-E^{-1}}\right] \varepsilon^0 - \frac{h^3}{12} \frac{1}{1-E^{-1}} (f''(\eta_m^1)). \quad (3.20)$$

By virtue of (3.20) and our assumptions on $f(t)$, we may write

$$(p_m^n) = \frac{h}{2} \frac{1+E^{-1}}{1-E^{-1}}(p_m^{n-1}) - \frac{h^3}{12} \frac{1}{1-E^{-1}} (f^{(3-n)}(\eta_m^n)), \; n=2,3,4,\ldots, \quad (3.21)$$

where $$f^{(-k)}(t) = \int_0^t \int_0^{t^k} \cdots \int_0^{t^2} f(t^1)\,dt^1 \ldots dt^{k-1}dt^k$$

and $f^{(k)}(t) = d^kf/dt^k$ for positive values of k, while $f^{(0)}(t) = f(t)$. Successive reduction, using (3.20) and (3.21) leads to

$$(p_m^n) = \left[\frac{h}{2}\,\frac{1+E^{-1}}{1-E^{-1}}\right]^n F(E^{-1})\varepsilon^0 - \left(\frac{h}{2}\right)^n f_0\,\frac{(1+E^{-1})^{n-1}}{(1-E^{-1})^n}\,\varepsilon^0 \tag{3.22}$$

$$- \frac{1}{12}\sum_{j=0}^{n-1} \frac{h^{j+3}}{2^j}\,\frac{(1+E^{-1})^j}{(1-E^{-1})^{j+1}}\left(f^{(j+3-n)}(\eta_m^{n-j})\right).$$

This equation is of the form

$$(p_m^n) = P^{-n}(E^{-1})F(E^{-1})\varepsilon^0 - C_n(E^{-1})f_0\,\varepsilon^0 - (\varepsilon_m^n) \tag{3.23}$$

where P^{-n}, C_n, (ε_m^n) may be designated respectively the integrating operator of order n, the correction operator of order n, and the error vector of order n.

Type II Operators.--If we let

$$\begin{aligned} g^n(t) &= t^{n-1}/(n-1)! && \text{for } 0 \le t \\ &= 0 && \text{for } t < 0 \end{aligned} \tag{3.24}$$

then (3.11) may be written

$$b_{m,k}^n = \int_{t_{k-1}}^{t} f(s)g^n(t_m-s)\,ds. \tag{3.25}$$

For a fixed value of m, the function $f(t)g^n(mh-t)$ satisfies the requirements for representation by means of an equation of the form (3.5). This representation is

$$f(t)g^n(mh-t) = \frac{1}{h}(f_k g^n_{m-k} - f_{k-1}g^n_{m-k+1})(t-t_{k-1}) \tag{3.26}$$

$$+ f_{k-1}g^n_{m-k+1} + \frac{1}{2}(t-t_k)(t-t_{k-1})\{\frac{d^2}{dt^2}[f(t)g^n(mh-t)]\}_{t = \xi_k}$$

where $g^n_m = g^n(mh)$. From this and (3.25) we obtain

$$b^n_{m,k} = \frac{h}{2}(f_k g^n_{m-k} + f_{k-1}g^n_{m-k+1}) - \frac{h^3}{12}\{\frac{d^2}{dt^2}[f(t)g^n(mh-t)]\}_{t = \eta_k}. \tag{3.27}$$

Substituting this result into (3.12) yields, after simplification,

$$(p^n_m) = (h\sum_{k=0}^{m} f_k g^n_{m-k}) - \frac{h}{2}(f_0 g^n_m + f_m g^n_0) - \frac{h^3}{12}\sum_{k=0}^{m}\{\frac{d^2}{dt^2}[f(t)g^n(mh-t)]\}_{t = \eta_k}. \tag{3.28}$$

The last equation may be written as

$$(p^n_m) = h\left[G_n(E^{-1}) - \frac{g^n_0}{2}\right]F(E^{-1})\ \varepsilon^0 \ - \frac{h}{2}\,G_n(E^{-1})f_0\ \varepsilon^0 \tag{3.29}$$

$$- \frac{h^3}{12(n-1)!}(\sum_{k=0}^{m}[f''(\eta_k)(mh-\eta_k)^{n-1} - 2(n-1)f'(\eta_k)(mh-\eta_k)^{n-2}$$

$$+ (n-1)(n-2)f(\eta_k)(mh-\eta_k)^{n-3}])$$

which is of the form (3.23).

Except for the case $n = 1$, $g^n_0 = 0$; for the exceptional case we note that we have $g^1_0 = g\ (t) = 1$ on the interval $0 \leq t \leq mh$. This yields a first order operator that is identical with the first order Type I and Type IV operators. Previous derivations have effectively taken $g_0^1 = 0$, but for $n > 1$, the above integrating operators coincide with Type II operators derived by other methods.

From (1.39) and (3.24) it is determined that $G_n(E^{-1})$ has the following expression in terms of Lawden's Ψ-functions (supra, p. 15):

$$G_1(E^{-1}) = \frac{1}{1-E^{-1}}; \tag{3.30}$$

$$G_n(E^{-1}) = \frac{h^{n-1}}{(n-1)!}\,\Psi_{n-1}(E^{-1}),\ n = 2,3,4,\dots\ . \tag{3.31}$$

Type IV Operators.--The derivation of these operators is similar to that of the Type II operators, except that rather than replacing the entire integrand of (3.11) by a piecewise linear approximation, only the function $f(t)$ itself is so approximated. Thus, replacing $f(s)$ in (3.11) by its equivalent form (3.5) and applying (3.6) results in

$$b_{mk}^n = \frac{1}{(n-1)!}\left[\frac{(f_k-f_{k-1})}{h}\int_{t_{k-1}}^{t_k}(t_m-s)^{n-1}(s-t_{k-1})ds \right. \tag{3.32}$$

$$\left. + f_{k-1}\int_{t_{k-1}}^{t_k}(t_m-s)^{n-1}\,ds + \frac{f''(\eta_k)}{2}\int_{t_{k-1}}^{t_k}(t_m-s)^{n-1}(s-t_k)(s-t_{k-1})ds\right].$$

Now letting

$$A_j^n(m,k) = \int_{t_{k-1}}^{t_k} s^j(t_m-s)^n ds,\quad n,\ j = 0,1,2,\dots, \tag{3.33}$$

we find that

$$A_0^{n-1}(m,k) = \frac{h^n}{n}[(m-k+1)^n - (m-k)^n]; \tag{3.34}$$

$$A_1^{n-1}(m,k) = \frac{h^{n+1}}{n}[(k-1)(m-k+1)^n - k(m-k)^n] + \frac{1}{n} A_0^n(m,k); \tag{3.35}$$

$$A_2^{n-1}(m,k) = \frac{h^{n+1}}{n}[(k-1)^2(m-k+1)^n - k^2(m-k)^n] + \frac{2}{n} A_1^n(m,k). \tag{3.36}$$

Substituting these values in (3.32) and simplifying the result gives

$$\begin{aligned} b_{mk}^n = \frac{h^n}{(n+1)!}\{&(f_k - f_{k-1})[(m-k+1)^{n+1} - (m-k)^{n+1}] \\ &+ (n+1)[f_{k-1}(m-k+1)^n - f_k(m-k)^n]\} \\ &- \frac{h^{n+2}}{2(n+2)!} f''(\eta_k)\{(n+2)[(m-k)^{n+1} + (m-k+1)^{n+1}] \\ &- 2[(m-k+1)^{n+2} - (m-k)^{n+2}]\}. \end{aligned} \tag{3.37}$$

Substituting this result in (3.12) gives

$$\begin{aligned} (p_m^n) = {} & \frac{h^n}{(n+1)!}\left(\sum_{k=0}^{m} (f_k - f_{k-1})[(m-k+1)^{n+1} - (m-k)^{n+1}]\right) \\ & - \frac{h^n f_0}{(n+1)!}((m+1)^{n+1} - m^{n+1} - (n+1)m^n) \\ & - \frac{h^{n+2}}{2(n+2)!}\left(\sum_{k=0}^{m} f''(\eta_k)\{(n+2)[(m-k)^{n+1} + (m-k+1)^{n+1}] \right. \\ & \left. - 2[(m-k+1)^{n+2} - (m-k)^{n+2}]\}\right). \end{aligned} \tag{3.38}$$

The (m+1)th component of the first vector on the right hand side of (3.38) is the coefficient of E^{-m} in the Cauchy product

$$\sum_{r=0}^{\infty}(f_r-f_{r-1})E^{-r}.\sum_{r=0}^{\infty}[(r+1)^{n+1}-r^{n+1}]E^{-r} \tag{3.39}$$

while the (m+1)th component of the third vector is the coefficient of E^{-m} in the Cauchy product

$$\sum_{r=0}^{\infty} f''(\eta_r)E^{-r}.\sum_{r=0}^{\infty}\{(n+2)[r^{n+1}+(r+1)^{n+1}] \tag{3.40}$$

$$-2[(r+1)^{n+2}-r^{n+2}]\}E^{-r}.$$

Taking (3.39) and (3.40) into account and performing the necessary expansions, we find that (3.38) becomes

$$(p_m^n)=\frac{h^n}{(n+1)!}(1-E^{-1})\{\sum_{r=0}^{\infty}[(n+1;1)r^n+\ldots+(n+1;k)r^{n-k+1}+\ldots+1]E^{-r}\}(f_m) \tag{3.41}$$

$$-\frac{h^nf_0}{(n+1)!}((n+1;2)m^{n-1}+\ldots+(n+1;k)m^{n-k+1}+\ldots+1)$$

$$-\frac{h^{n+2}}{2(n+2)!}\sum_{r=0}^{\infty}\{(n+2)[2r^{n+1}+(n+1;1)r^n+\ldots+(n+1;k)r^{n-k+1}+\ldots+1]$$

$$-2[(n+2;1)r^{n+1}+\ldots+(n+2;k)r^{n-k+1}+\ldots+1]\}E^{-r}.(f''(\eta_m)).$$

Lawden [7] establishes the following relation:

$$(z^{-1}-1)\,\Psi_n(z)=(n;1)\,\Psi_{n-1}(z)+(n;2)\,\Psi_{n-2}(z)+\ldots+\Psi_0(z)+1.$$

A relation more convenient for our purposes is obtained by replacing z by z^{-1} and n by n+1 in this equation:

$$\frac{(1-z^{-1})}{z^{-1}}\,\Psi_{n+1}(z^{-1})=(n+1;1)\,\Psi_n(z^{-1})+(n+1;2)\,\Psi_{n-1}(z^{-1})+\ldots$$

$$+(n+1;n)\,\Psi_1(z^{-1})+\Psi_0(z^{-1})+1.$$

Finally, applying this and (1.39) to (3.41) we have the desired result, of the form (3.23),

$$
\begin{aligned}
(p_m^n) = {} & \frac{h^n}{(n+1)!}\,\frac{(1-E^{-1})^2}{E^{-1}}\,\Psi_{n+1}(E^{-1})(f_m) \\
& - \frac{h^n f_0}{(n+1)!}\left[\frac{1-E^{-1}}{E^{-1}}\,\Psi_{n+1}(E^{-1}) - (n+1)\,\Psi_n(E^{-1})\right]\varepsilon^0 \\
& - \frac{h^{n+2}}{2(n+2)!}\left[(n+2)\frac{1+E^{-1}}{E^{-1}}\,\Psi_{n+1}(E^{-1}) - 2\,\frac{1-E^{-1}}{E^{-1}}\,\Psi_{n+2}(E^{-1})\right](f''(\eta_m)).
\end{aligned}
\tag{3.42}
$$

At this point, a comment on Type V operators is called for. In [41], it is stated erroneously that these operators are in a sense a variant of the Type II operators. Actually, the derivation of Type V operators from the present viewpoint is effected in almost precisely the same way as is that for the Type IV operators. The only difference is that, rather than using (3.5) in the representation of f(t), use is made instead of the mean value theorem for derivatives. Since the resulting operators are less accurate than those here considered, the details of the derivation are omitted, the foregoing remarks being included only to satisfy the reader's curiosity and to correct the erroneous statement of [41].

Bounds on the error for operators of types I, II and IV.-- Let $\overline{|h(t)|}^m$ denote the maximum value of $|h(t)|$ on the interval $[0,mh]$ and let $|H(E^{-1})(f_m)|_n$ denote the absolute value of the n-th component of the vector $H(E^{-1})(f_m)$. From (3.22) we obtain, for Type I operators,

$$|\varepsilon_m^n| \le \frac{1}{12} \sum_{j=0}^{n-1} \frac{h^{j+3}}{2^j} \overline{|f^{(j+3-n)}(t)|^m} \left| \frac{E^{-1}(1+E^{-1})^j}{(1-E^{-1})^{j+2}} \varepsilon^0 \right|_m \tag{3.43}$$

which yields

$$|\varepsilon_m^n| < \frac{1}{12} \sum_{j=0}^{n-1} \frac{h^{j+3}}{2^j} \overline{|f^{(j+3-n)}(t)|^m} \sum_{q=0}^{r} (j;q)(j+m-q;m-q-1) \tag{3.44}$$

where

$$\begin{aligned} r &= m-1 \quad \text{for } m-1 < j \\ &= j \quad \text{for } m-1 \ge j. \end{aligned}$$

Now

$$(j;q)(j+m-q;m-q-1) = \frac{1}{(j+1)q!(j-q)!} \underbrace{[(j-q)+m][(j-q)+(m-1)]\ldots[(j-q)+(m-j)]}_{j\,+\,1 \text{ factors}}; \tag{3.45}$$

hence, for fixed j and q, the interior sum of (3.44) is a polynomial of degree j+1 in m. Since mh is independent of h we conclude

$$|\varepsilon_m^n| \le h^2 T(h) \tag{3.46}$$

where T(h) is a polynomial in h whose coefficients are independent of h.

For Type IV operators we have from (3.42)

$$|\varepsilon_m^n| \le \frac{h^{n+2}}{2(n+2)!} \overline{|f''(t)|^m} \left| \left[(n+2)\frac{1+E^{-1}}{1-E^{-1}} \Psi_{n+1}(E^{-1}) - 2\,\Psi_{n+2}(E^{-1}) \right] \varepsilon^0 \right|_m . \tag{3.47}$$

In particular, by using (1.39) in conjunction with (3.47), we find that

$$|\varepsilon_m^n| \le \overline{|f''(t)|^m} \frac{h^{n+2}}{2(n+2)!} [(n+2)m^{n+1} - 2m^{n+2} + 2(n+2) \sum_{j=1}^{m-1} j^{n+1}]. \tag{3.48}$$

By use of a well-known expression for the summation appearing in (3.48), (cf. [42], p. 137), (3.48) reduces to, for m > 1,

$$|\varepsilon_m^n| \le \frac{\overline{|f''(t)|^m}}{(n+2)!} [(n+2;2)B_2(mh)^n h^2 + \ldots + (n+2;n+1)B_{n+1}(mh)h^{n+1}] \tag{3.49}$$

where B_i is the i-th first order Bernoulli number [42], so that for this type of operator also, the error magnitude is of the order $O(h^2)$.

Finally, a similar analysis of the error vector of (3.29) yields for Type II operators with $n > 3$

$$\begin{aligned}|\varepsilon^n_m| \leq \frac{h^3}{12(n-1)!} \Bigg[& \frac{h^{n-1}}{n} \overline{|f''(t)|}^m \sum_{j=1}^{n} (n;j) B_{n-j} (m+1)^j \qquad (3.50) \\ & + 2h^{n-2} \overline{|f'(t)|}^m \sum_{j=1}^{n-1} (n-1;j) B_{n-j-1} (m+1)^j \\ & + (n-1) h^{n-3} \overline{|f(t)|}^m \sum_{j=1}^{n-2} (n-2;j) B_{n-j-2} (m+1)^j \Bigg]\end{aligned}$$

from which it is clear that the magnitude of the error is again of the order $O(h^2)$. The same order of magnitude can be obtained for $n = 1,2,3$.

Hence, if we judge the error only by its order of magnitude, we see that the three types of operators possess the same degree of accuracy. It is well to note that if $f_0 \neq 0$ and we use only $P^{-n}(E^{-1})(f_m)$ to approximate (p^n_m), then the error, for all three types of operators, is of the order $O(h)$. This is verified by observing that this is the order of magnitude of the m-th component of the vector $f_0 C_n(E^{-1}) \varepsilon^0$.

A deeper analysis of the error does not appear easy. However, several arguments may be advanced for preferring the use of Type IV operators in the applications, the most compelling of which is that the error vector contains the least number of independent quantities of the three types of operators we have studied.

Approximate solution of linear ordinary differential equations.-- We shall apply the results of the preceding section to the approximate solution of linear ordinary differential equations of the form

$$\sum_{j=0}^{n} \frac{d^j}{dt^j}(a_{n-j}y) = f(t) \tag{3.51}$$

where the a_i, $i = 0,1,\ldots,n$, are functions only of the independent (real) variable t. We assume that a_{n-j} is continuous on every finite interval $[0,t]$ and j times differentiable on the interior of the interval. It is assumed further that f(t) is an integrable but otherwise arbitrary function on the same interval. Finally, we assume $a_0 \neq 0$ on the interval.

The equation of (3.51) is not as restrictive as it appears at first glance, for by use of the identity

$$\frac{d^k}{dt^k}(xy) = \sum_{j=0}^{k}(k;j)\frac{d^{k-j}x}{dt^{k-j}} \cdot \frac{d^jy}{dt^j}, \tag{3.52}$$

where $(k;j) = k!/[j!(k-j)!]$, (3.51) may be put in the form

$$\sum_{j=0}^{n} b_{n-j}\frac{d^jy}{dt^j} = f(t) \tag{3.53}$$

with

$$b_j = \sum_{k=0}^{j}(n-k;n-j)\frac{d^{j-k}}{dt^{j-k}}a_k, \quad j = 0,1,\ldots,n. \tag{3.54}$$

With the additional assumption that the b_j are n-j times differentiable, an inductive argument from (3.54) yields the reciprocal relations

$$a_j = \sum_{k=0}^{j}(-1)^{j-k}(n-k;n-j)\frac{d^{j-k}}{dt^{j-k}}b_k, \quad j = 0,1,2,\ldots,n.$$

The proof is relegated to the Appendix. Thus the subsequent analysis applies as well to equations of the form (3.53) with the just-noted restriction on the b_j.

Integration of (3.51) between the limits 0, t,--n times repeated--results in

$$\sum_{j=0}^{n} D^{-j}(a_j y) = \varphi(t) \tag{3.55}$$

where

$$D^{-k}y = \int_0^t \int_0^{t_k} \dots \int_0^{t_2} y(t_1)dt_1 \dots dt_{k-1}dt_k, \quad k = 1,2,3,\dots; \tag{3.56}$$

$$D^{-0}y(t) = y(t)$$

$$F(t) = D^{-n}f(t);$$

$$\varphi(t) = F(t) + \sum_{j=0}^{n-1} \sum_{k=0}^{j} \frac{K_{j-k}^{k} t^j}{j!}; \tag{3.57}$$

$$K_j^k = \left[\frac{d^j}{dt^j}(a_k y)\right]_{t=0}, \quad j, k = 0,1,\dots,n-1. \tag{3.58}$$

In (3.55), we put t = mh and use (3.23) to write the corresponding vector-matrix equation:

$$\sum_{k=0}^{n} \{P^{-k}(E^{-1})(a_{km}y_m) - a_{k0}y_0C_k(E^{-1})\varepsilon^0 - (\varepsilon_m^k)\} = (\varphi_m). \tag{3.59}$$

Here we have expanded the notation used earlier to include

$$P^{-0}(E^{-1}) = 1;$$

$$C_0(E^{-1}) = 0;$$

$$(\varepsilon_m^0) = (0,0,\dots).$$

We define the vector $(\bar{y}_m)$ by means of the equation

$$\sum_{k=0}^{n} \{P^{-k}(E^{-1})(a_{km}\bar{y}_m) - a_{k0}y_0C_k(E^{-1})\ \varepsilon^0\} = (\varphi_m); \tag{3.60}$$

then, putting $\delta_m = y_m - \bar{y}_m$, we obtain from (3.59) and (3.60)

$$\sum_{k=0}^{n} \{P^{-k}(E^{-1})(a_{km}\delta_m)\} = \sum_{k=0}^{n} (\varepsilon_m^k). \tag{3.61}$$

Now let

$$P^{-k}(E^{-1}) = \sum_{r=0}^{\infty} p_r^k\, E^{-r};$$

then the corresponding d.i.m. is (ρ_{ij}^k) where

$$\rho_{ij}^k = p_{i-j}^k.$$

Using this, we write the vector-matrix equation equivalent to (3.61):

$$\sum_{k=0}^{n} (\rho_{ij}^k)(a_{km}\delta_m) = (\sum_{k=0}^{n} \varepsilon_m^k). \tag{3.62}$$

Carrying out the indicated multiplication in (3.62) results in

$$\sum_{k=0}^{n} (\sum_{j=0}^{m} p_{m-j}^k a_{kj}\delta_j) = (\sum_{k=0}^{n} \varepsilon_m^k), \tag{3.63}$$

which in turn yields

$$(\sum_{j=0}^{m} \sum_{k=0}^{n} p_{m-j}^k a_{kj}\delta_j) = (\sum_{k=0}^{n} \varepsilon_m^k). \tag{3.64}$$

Now define λ_{ij}^n as

$$\lambda_{ij}^n = \sum_{k=0}^{n} p_{i-j}^k a_{k,j-1},\quad i,j = 1,2,3,\dots\ . \tag{3.65}$$

Using (3.65), (3.64) may be written as

$$\left(\sum_{j=0}^{m} \lambda^{n}_{m+1,j+1}\delta_j\right) = \left(\sum_{k=0}^{n} \varepsilon^{k}_{m}\right) \tag{3.66}$$

or, equivalently, as

$$(\lambda^{n}_{ij})(\delta_m) = \left(\sum_{k=0}^{n} \varepsilon^{k}_{m}\right). \tag{3.67}$$

From the foregoing, it is clear that if we put

$$(\mu_m) = (\varphi_m) + y_0 \sum_{k=0}^{n} a_{k0}C_k(E^{-1})\ \varepsilon^0 \tag{3.68}$$

then (3.60) may be written as

$$(\lambda^{n}_{ij})(\bar{y}_m) = (\mu_m). \tag{3.69}$$

The formulation is now complete. In the following sections we shall first use (3.66) to demonstrate that $\bar{y}_m$ approximates y_m with the accuracy $O(h^2)$ and then solve (3.69) for $\bar{y}_m$.

Accuracy of the approximate solution.--From (3.66) it is seen that $\delta_0 = 0$. For $m \geq 1$, we readily obtain

$$|\delta_m| \leq \frac{1}{|\lambda^{n}_{m+1,m+1}|}\left\{\sum_{k=1}^{n} |\varepsilon^{k}_{m}| + \alpha_{m-1}\sum_{j=1}^{m-1} |\lambda^{n}_{m+1,m+1}|\right\},\ m = 1,2,3,\ldots \tag{3.70}$$

$$\leq \frac{1}{|a_{0m}| - \sum_{k=1}^{n} |p^{k}_{0}||a_{km}|}\left\{\sum_{k=1}^{n} |\varepsilon_m| + \alpha_{m-1}\sum_{j=1}^{m-1}\sum_{k=0}^{n} |p^{k}_{m-j}|\cdot|a_{kj}|\right\} \tag{3.71}$$

where $\alpha_m = \max\limits_{1\leq r\leq m} |\delta_r|$.

By our initial assumptions about the $a_j(t)$, $j = 0,1,\dots,n$, there exist positive numbers L_0, U_0, $U_1,\dots,U_n$ such that

$$\left.\begin{array}{l} L_0 \le |a_0(t)| \le U_0 \\ |a_j(t)| \le U_j \end{array}\right\} \text{for } 0 \le t \le mh.$$

Using these bounds with (3.71) yields

$$|\delta_m| \le \frac{1}{L_0} \cdot \frac{1}{1 - \frac{1}{L_0} \sum_{k=1}^{n} U_k |p_0^k|} \{ \sum_{k=1}^{n} |\varepsilon_m^k| + \alpha_{m-1} \sum_{j=1}^{m-1} \sum_{k=0}^{n} U_k |p_{m-j}^k| \}. \tag{3.72}$$

Now suppose that mh is fixed and that h_1 is a fixed constant such that $0 < h \le h_1$. Then any of (3.46), (3.49), (3.50) imply that

$$\sum_{k=1}^{m} |\varepsilon_m^k| \le Kh^2$$

where K is independent of h.

If, for a sufficiently small value of h_1, $1 - \frac{1}{L_0} \sum_{k=1}^{n} U_k |p_0^k|$ is uniformly bounded away from zero on $[0,h_1]$ and if $\sum_{j=1}^{m-1} \sum_{k=0}^{n} U_k |p_{m-j}^k|$ is uniformly bounded above on the same interval, then the condition

$$\alpha_{m-1} \le Lh^2 \tag{3.73}$$

implies, through (3.72), that

$$|\delta_m| \le Mh^2 \tag{3.74}$$

where L,M are constants. Moreover, since

$$\alpha_m = \max\,[|\delta_m|, \alpha_{m-1}],$$

(3.73) and (3.74) imply that

$$\alpha_m = O(h^2). \tag{3.75}$$

From (3.72) we have

$$\alpha_1 = |\delta_1| \leq \frac{1}{L_0} \frac{1}{1 - \frac{1}{L_0} \sum_{k=1}^{n} U_k |p_0^k|} \sum_{k=1}^{n} |\epsilon_1^k| \leq Rh^2$$

and from this we infer by mathematical induction that (3.75) is true for all positive integral values of m, contingent of course upon the boundedness conjectured previously. This boundedness is established in the Appendix.

We conclude that, for each of the operators considered herein, $\bar{y}_m$ approximates y_m to the order $O(h^2)$ for positive integral values of m. As noted previously, $\bar{y}_0 = y_0$. Parenthetically it should be observed that an equation similar to (3.67), and used in the same connection, appears to be treated incorrectly by Wasow [40].

The solution of (3.69).--The solution of this equation is of course just

$$(\bar{y}_m) = (\lambda_{ij}^{n})^{-1}(\mu_m) = (\lambda_{ij}^{-n})(\mu_m) \tag{3.76}$$

when the inverse exists. For this existence, the following condition is sufficient (supra, p. 7):

$$\lambda_{ii}^{n} = \sum_{k=0}^{n} p_0^k a_{k,i-1} \neq 0, \quad i = 1,2,3,\dots . \tag{3.77}$$

In the event that (3.77) is violated for every value of i, then a pseudo-reciprocal may still exist, the sufficient condition for this existence being an analogue of (3.77).

With the assumption that (3.77) is satisfied, the determination of (λ_{ij}^{-n}) is obtained from (1.10) as

$$\lambda_{ij}^{-n} = (-1)^{i-j} \det \omega_{ij}^{n} / \prod_{l=j}^{i} \lambda_{ll}^{n} \tag{3.78}$$

where

$$\det \omega_{ii}^{n} = 1; \tag{3.79}$$

$$\det \omega_{ij}^{n} = \begin{vmatrix} \lambda_{j+1,j} & \lambda_{j+1,j+1} & 0 & 0 \ldots 0 & 0 \\ \lambda_{j+2,j} & \lambda_{j+2,j+1} & \lambda_{j+2,j+2} & 0 \ldots 0 & 0 \\ \cdot & \cdot & \cdot & \cdot\quad\cdot\quad\cdot & \cdot \\ \lambda_{i-1,j} & \lambda_{i-1,j+1} & \cdot\;\cdot\;\cdot & \lambda_{i-1,i-2} & \lambda_{i-1,i-1} \\ \lambda_{ij} & \lambda_{i,j+1} & \cdot\;\cdot\;\cdot & \lambda_{i,i-2} & \lambda_{i,i-1} \end{vmatrix}, j=1,2,\ldots,i-1. \tag{3.80}$$

Alternatively, from (1.11),

$$\lambda_{ii}^{-n} = 1/\lambda_{ii}^{n}; \tag{3.81}$$

$$\lambda_{ij}^{-n} = -[\lambda_{ij}^{n}\lambda_{ii}^{-n} + \lambda_{i-1,j}^{n}\lambda_{i,i-1}^{-n} + \ldots + \lambda_{j+1,j}^{n}\lambda_{i,j+1}^{-n}]/\lambda_{jj}^{n}, \tag{3.82}$$

$$j = 1,2,\ldots,i-1.$$

In the event that the coefficients $a_j(t)$ of (3.51) are constants--call them a_j--then the equivalent forms of (3.78), (3.79) and (3.80) for this case can be obtained directly from (1.12) and (1.13). The displacement transformation corresponding to the resulting d.i.m. for the solution $\bar{y}_m$ is readily obtained by the methods of the first chapter as

$$\left[1 + \sum_{j=1}^{n} a_j P^{-j}(E^{-1})\right](\bar{y}_m) = \left\{ \Phi(E^{-1}) + \sum_{s=1}^{n-1} \left[\sum_{r=0}^{s} a_r y^{(s-r)}(0) \right] \frac{h^s \Psi_s(E^{-1})}{s!} + y_0 \left[\frac{1}{1-E^{-1}} + \sum_{j=1}^{n} a_j C_j(E^{-1}) \right] \right\} \varepsilon^0 \qquad (3.83)$$

where $\Phi(E^{-1})\varepsilon^0 = (f_m)$. This representation of the approximate solution $\bar{y}_m$ is easier to use than the corresponding d.i.m., for it is only required to operate on (3.83) by the inverse (or pseudoinverse) of $1 + \sum_{j=1}^{n} a_j P^{-j}(E^{-1})$. Since this operator is rational, as is $\Phi(E^{-1})$ in most applications, the indicated operation amounts to nothing more than long division of one polynomial by another.

Remarks.--It has previously been remarked that Type IV operators are to be preferred for approximate quadrature by virtue of the simplicity of the associated bound on the error. This preference is retained in the application to constant coefficient differential equations for the same reason.

Our analysis has not precluded the use of any of the three types of operators in the variable coefficient case; however, Type I operators may be dismissed on a practical basis, since the general form of the coefficients p^k_{i-j} is too cumbersome for computational purposes when compared to that of the other types. It is worthwhile to exhibit here the general forms of the p^k_{i-j} for Type II and Type IV operators for $k \geq 1$:

$$\left.\begin{aligned} p_0^1 &= \frac{h}{2} \\ p_{i-j}^1 &= h, \ i > j \\ p_{i-j}^k &= \frac{h^k}{(k-1)!}(i-j)^{k-1}, \ i \geq j, \ k > 1 \end{aligned}\right\} \qquad \text{(Type II);} \tag{3.84}$$

$$\left.\begin{aligned} p_0^k &= \frac{h^k}{(k+1)!} \\ p_{i-j}^k &= \frac{h^k}{(k+1)!}\left[(i-j+1)^{k+1} - 2(i-j)^{k+1} + (i-j-1)^{k+1}\right], \ i > j \end{aligned}\right\} \text{(Type IV).} \tag{3.85}$$

Of course, for both types we have

$$p_{i-j}^0 = \delta_{ij},$$

the Kronecker delta.

It thus appears that in the variable coefficient case the computational convenience of the Type II operators may be an overriding advantage. Perhaps even more important is the fact that, for Type II operators, condition (3.77) reduces to

$$\lambda_{ii}^n = a_{0,i-1} + \frac{h}{2} a_{1,i-1} \neq 0, \ i = 1,2,3,\ldots,$$

which is considerably simpler than the corresponding expression of the condition for Type IV operators.

COMPENDIUM

In this dissertation an attempt has been made to establish on a rigorous basis two fields of current interest in applied mathematics and engineering. As a consequence, many of the important mathematical expressions that have appeared in the foregoing pages have already appeared--in varying degrees of generality--in the work of earlier investigators in these fields. It is not amiss therefore to enumerate, for the benefit of the reader who is not a cognoscente of these fields, what the author deems to be the primary contributions of this dissertation to the subject fields.

A class of linear transformations on a sequence space--the linear filters--has been discussed and shown to constitute a suitable--even a "natural"--mathematical representation for the class of linear sampled-data systems. This representation on the one hand justifies the use of infinite matrices--implicit in the work of Friedland [12][7]--in the analysis of sampled-data systems. On the other hand, the introduction of the displacement transformation as an alternate representation of a constant linear filter provides the necessary justification for the use of the z-transform in the study of time-invariant sampled-data systems. Moreover, the complete equivalence, in this case, of the matrix and transformation techniques has been established.

---

[7] Since the completion of the penultimate draft of this dissertation, a more recent paper by Friedland--"Theory of time-varying sampled-data systems," Tech. Rep. T-19/B, Columbia University Electronic Research Laboratories, April, 1957--has come to the author's attention. In this paper Friedland indicates an awareness, not evident in [12], of the applicability of infinite matrix theory. In particular, he utilizes the Kojima-Schur theorem for stability considerations. However, being more concerned with the "w-transform" and with applications, he has not pursued the topic to the depth and generality of the present work.

Two mappings--one an isomorphism and the other a homomorphism--of the class of linear sampled-data systems on the class of linear filters have been studied. In the class of constant linear filters, the isomorphic mapping leads to the counterpart of Jury's "modified" z-transform [14] while the homomorphic mapping leads to the counterpart of the "classical" z-transform; one conclusion to be drawn is that Jury's transform is not a modification of the z-transform but, on the contrary, the z-transform is to be considered a special case of Jury's transform. In the class of variable linear filters the homomorphic mapping leads to the matrix formulation which is the counterpart of Friedland's [12]. The isomorphic mapping leads to a more general formulation--of which Friedland's is a special case--which apparently has not been previously derived although the possibility of such a formulation has been discussed by Friedland in a recent paper (see footnote 7).

Finally, with regard to sampled-data systems, the importance of the use of Lawden's Ψ-functions [7] for canonical representation of the time-invariant sampled-data systems has been indicated. This point seems to have been overlooked by American investigators and to have been exploited to only a limited extent by British investigators.

The remainder of this work has consisted of a thorough analytical study of the method of numerical quadrature conceived by Tustin [18], [19] and extended in various ways by others. Use of the dual matrix-transformation techniques developed in the first chapter has permitted putting Tustin's ideas and their extensions on a rigorous foundation involving only elementary concepts from the calculus. The development has at the same time permitted retention of the operational form characteristic of Tustin's methods.

The importance of a correction for non-zero values of the integrand has been put in evidence together with a general form for the correction term which is believed to be new, as is the expression for the general form of Madwed's integrating operators. Bounds on the error for this method of numerical quadrature have been derived in a form essentially different from those given by earlier investigators. Finally, a new algorithm for the application of the method to the numerical solution of linear ordinary differential equations has been derived together with an estimate of accuracy.

LIST OF REFERENCES

1. R. G. Cooke, *Infinite Matrices and Sequence Spaces*, Macmillan, London, 1950.

2. G. Matthews, *A class of infinite matrices*, Proc. Koninklijke Nederlandse Akademie van Wetenschappen, ser. A, vol. 51, no. 4 (1958), pp. 457-59.

3. E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*, Colloq. Pub. 31, Amer. Math. Soc., Providence, 1957.

4. R. R. Stoll, *Linear Algebra and Matrix Theory*, McGraw-Hill, New York, 1952.

5. J. Hagen, *On division of series*, Amer. J. Math., vol. 5 (1883), p. 236.

6. W. Hurewicz, *Theory of Servomechanisms*, ed. H. M. James, N. B. Nichols, and R. S. Phillips, McGraw-Hill, New York, 1947, chap. 5.

7. D. F. Lawden, *The function $\sum_{n=1}^{\infty} n^r z^n$ and associated polynomials*, Proc. Cambridge Philos. Soc., vol. 47 (1951), pp. 309-14.

8. G. J. Murphy and R. D. Ormsby, *A survey of techniques for the analysis of sampled-data control systems*, Trans. Inst. Radio Engrs., PGAC-2 (Feb. 1957), pp. 79-90.

9. H. Freeman and O. Lowenschuss, *Bibliography of sampled-data control systems and z-transform applications*, Trans. Inst. Radio Engrs., PGAC-4 (Mar. 1958), pp. 28-30.

10. P. R. Stromer, *A selective bibliography on sampled-data systems*, Trans. Inst. Radio Engrs., PGAC-6 (Dec. 1958), pp. 112-14.

11. L. A. MacColl, *Fundamental Theory of Servomechanisms*, D. Van Nostrand, New York, 1945.

12. B. Friedland, *A technique for the analysis of time-varying sampled-data systems*, Trans. Amer. Inst. Elec. Engrs., vol. 75, pt. 2 (Jan. 1957), pp. 407-14.

13. L. A. Zadeh, *A general theory of linear signal transmission systems*, J. Franklin Inst., vol. 253 (Apr. 1952), pp. 293-312.

14. E. I. Jury, *Synthesis and critical study of sampled-data control systems*, Trans. Amer. Inst. Elec. Engrs., vol. 75, pt. 2 (July 1956), pp. 141-51.

15. J. R. Ragazzini, Digital computers in feedback systems, Inst. Radio Engrs. National Convention Record, pt. 4 (1957), pp. 33-42.

16. H. M. James and P. R. Weiss, Theory of Servomechanisms, ed. H. M. James, N. B. Nichols, and R. S. Phillips, McGraw-Hill, New York, 1947, chap. 2.

17. B. Friedland, Transformation techniques for time-varying sampled-data systems, Columbia University Electronics Research Laboratories Tech. Rep. T-13/B (1956).

18. A. Tustin, A method of analyzing the behavior of linear systems in terms of time series, J. Inst. Elec. Eng., vol. 94, pt. 2a (1947), pp. 130-42.

19. _______, A method of analyzing the effect of certain kinds of nonlinearity in closed-cycle control systems, J. Inst. Elec. Eng., vol. 94, pt. 2a (1947), pp. 152-60.

20. A. Madwed, Number series method of solving linear and nonlinear differential equations, MIT Inst. Lab. Rep. 6445-T-26 (1950).

21. F. W. Bubb, Cardinal spectrum analysis of physical systems, Wright Air Development Center, Air Research and Development Command, Air Force Tech. Rep. 6300 (1951).

22. _______, A new linear operational calculus, Wright Air Development Center, Air Research and Development Command, Air Force Tech. Rep. 6581 (1951).

23. _______, Linear noise smoothing and predicting filters, Wright Air Development Center, Air Research and Development Command, Air Force Tech. Rep. 6586 (1951).

24. R. G. Piety, A linear operational calculus of empirical functions, Phillips Petroleum Co., Research Div., Rep. 106-12-51R (1951).

25. B. M. Brown, Application of finite difference operators to linear systems, Automatic and Manual Control, ed. A. Tustin, Academic Press, New York, 1952, pp. 409-18.

26. J. G. Truxal, Numerical analysis for network design, Trans. Inst. Radio Engrs., vol. CT-1, no. 3 (1954), pp. 49-60.

27. J. R. Ragazzini and A. R. Bergen, A mathematical technique for the analysis of linear systems, Proc. Inst. Radio Engrs., vol 42 (1954), pp. 1645-51.

28. S. Amarel, *A computational method for treating networks containing nonlinear elements*, Columbia University Electronics Research Laboratories Tech. Rep. T-3/c (1954).

29. _______, *An operational method for the treatment of transients in wave-form circuits (linear and non-linear)*, Columbia University Electronics Research Laboratories Tech. Rep. T-5/c (1955).

30. _______, L. Lambert, and J. Millman, *A transformation calculus for obtaining approximate solutions to linear integro-differential equations with constant coefficients*, Columbia University Electronics Research Laboratories Tech. Rep. T-7/c (1955).

31. A. J. O. Cruickshank, *A note on time series and the use of jump functions in approximate analysis*, Proc. Inst. Elec. Engrs., vol. 102, pt. C (1955).

32. W. Wasow, *Approximate integration of ordinary differential equations by means of Laplace transformation*, Polytechnic Institute of Brooklyn, Microwave Research Rep. R-504-56, PIB-434 (1956).

33. R. Boxer and S. Thaler, *A simplified method of solving linear and nonlinear systems*, Proc. Inst. Radio Engrs., vol. 44 (1956), pp. 89-101.

34. _______, *Extensions of numerical transform theory*, Rome Air Development Center, Tech. Rep. 56-115 (1956).

35. S. Thaler and R. Boxer, *An operational calculus for numerical analysis*, Inst. Radio Engrs. National Convention Record (1956).

36. _______, *Algebraic tables for a numerical transform calculus*, Rome Air Development Center, Tech. Note 56-281 (1956).

37. M. Cuenod, *Methode de calcul a l'aide de suites*, Science et Technique, Switzerland (1956).

38. R. Boxer, *A note on numerical transform calculus*, Proc. Inst. Radio Engrs., vol. 45 (1957), pp. 1401-06.

39. J. Kukel, *Sampling in linear and non-linear feedback control systems*, Inst. Radio Engrs. National Convention Record, pt 4 (1957).

40. W. Wasow, *Discrete approximations to the Laplace transformation*, Z. Angew. Math. Phys., vol. 8 (1957), pp. 401-17.

41. T. F. Bridgland, Jr., A note on numerical integrating operators, J. Soc. Indust. Appl. Math., vol. 6 (Sept. 1958), pp. 240-56.

42. L. M. Milne-Thomson, The Calculus of Finite Differences, Macmillan, London, 1951.

APPENDIX

Definition of "algebra".--The definition parallels that of [3]. A set X is an algebra over a field of scalars, Φ, if and only if the elements of X admit the three operations of addition, multiplication and multiplication by a scalar, subject to the following conditions:

i) X is an Abelian group under addition

ii) for every $\alpha \epsilon \Phi$ and every $x \epsilon X$, there is a uniquely defined product $\alpha x = x\alpha$ in X (multiplication by a scalar)

$$\left.\begin{aligned} &\text{iii)}\quad (\alpha+\beta)x = \alpha x + \beta x \\ &\text{iv)}\quad \alpha(x+y) = \alpha x + \alpha y \\ &\text{v)}\quad \alpha(\beta x) = (\alpha\beta)x \\ &\text{vi)}\quad 1\cdot x = x \end{aligned}\right\} \text{ where } x,y \epsilon X \text{ and } 1,\alpha,\beta \epsilon \Phi$$

vii) for every ordered pair $x,y \epsilon X$, there is a unique product $xy \epsilon X$

$$\left.\begin{aligned} &\text{viii)}\quad (xy)z = x(yz) \\ &\text{ix)}\quad x(y+z) = xy + xz;\ (x+y)z = xz + yz \\ &\text{x)}\quad \alpha x\beta y = \alpha\beta xy \end{aligned}\right\} \begin{aligned} &\text{where } x,y,z \epsilon X \\ &\text{and } \alpha,\beta \epsilon \Phi \end{aligned}$$

If in addition there exists an element e such that

xi) $ex = xe = x$ for every $x \epsilon X$

then X is called an algebra with unit element.

If, for every $x,y \epsilon X$,

xii) $xy = yx$

then X is called a commutative (Abelian) algebra

If X satisfies conditions i through vi only, it is called a linear vector space over ϕ.

Derivation of relations reciprocal to those of (3.54).--

Let us suppose that

$$a_j = \sum_{k=0}^{j} (-1)^{j-k}(n-k;n-j)\frac{d^{j-k}}{dt^{j-k}}\, b_k, \quad j = 0,1,\dots,m-1, \tag{A1.1}$$

where $m \leq n$. Then, letting $j = m$ in (3.54), and using (A1.1) we obtain

$$a_m = b_m - \sum_{k=0}^{m-1} \sum_{j=0}^{k} (-1)^{k-j}(n-k;n-m)(n-j;n-k)\frac{d^{m-j}}{dt^{m-j}}\, b_j. \tag{A1.2}$$

Now, it is trivial to show that

$$(n-k;n-m)(n-j;n-k) = (n-j;n-m)(m-j;m-k)$$

so that, assuming $(a;b) = 0$ for $a < b$, (A1.2) can be written

$$a_m = b_m - \sum_{j=0}^{m-1} \sum_{k=j}^{m-1} (-1)^{k-j}(n-j;n-m)(m-j;m-k)\frac{d^{m-j}}{dt^{m-j}}\, b_j \tag{A1.3}$$

or

$$a_m = b_m + \sum_{j=0}^{m-1} (-1)^{m-j}(n-j;n-m)\frac{d^{m-j}}{dt^{m-j}}\, b_j \Big\{ \sum_{k=j}^{m-1} (-1)^{m-k+1}(m-j;m-k)\Big\}. \tag{A1.4}$$

But the quantity in braces in (A1.4) has, by a well-known property of the binomial coefficients, the value one. From (3.54) we have

$$a_0 = b_0$$

so that with this and (A1.4) it follows by finite induction that (A1.1) is true for $j = 0,1,\dots,n$.

Oddly enough, the author has been unable to find this interesting and useful little theorem proven elsewhere.

Proof of boundedness of expressions contained in (3.72).--

From (3.84), (3.85) and the expression

$$p_0^k = \left(\frac{h}{2}\right)^k \tag{A1.5}$$

for Type I operators, it is clear that

$$1 - \frac{1}{L_0} \sum_{k=1}^{n} U_k |p_0^k| \tag{A1.6}$$

is uniformly bounded away from zero when $0 \leq h \leq h_1$ provided a sufficiently small value of h_1 is taken.

For the remaining expression,

$$\sum_{j=1}^{m-1} \sum_{k=1}^{n} U_k |p_{m-j}^k| \tag{A1.7}$$

the boundedness is demonstrated by showing that, for each of the three types of operators, (A1.7) is a polynomial in h on any interval $[0,h_1]$. The maximum of the polynomial on any such interval then serves as a uniform upper bound.

First, (A1.7) is replaced by its simpler equivalent

$$\sum_{j=1}^{m-1} \sum_{k=1}^{n} U_k |p_j^k| \tag{A1.8}$$

Using (3.84), (A1.8) becomes, for Type II operators,

$$\sum_{j=1}^{m-1} \{ U_1 h + \sum_{k=2}^{n} \frac{U_k h^k j^{k-1}}{(k-1)!} \} \tag{A1.9}$$

$$= U_1(mh) - U_1 h + \sum_{k=2}^{n} \frac{U_k h^k}{(k-1)!} \left(\sum_{j=1}^{m-1} j^{k-1} \right) \tag{A1.10}$$

$$= U_1(mh) - U_1 h + \sum_{k=2}^{n} \frac{U_k h^k}{k!} [m^k + (k;1)B_1 m^{k-1} + \ldots + (k;k-1)B_{k-1} m] \tag{A1.11}$$

$$= U_1(mh) - U_1 h + \sum_{k=2}^{n} \frac{U_k}{k!} \sum_{s=0}^{k-1} (k;s) B_s (mh)^{k-s} h^s . \tag{A1.12}$$

The transition from the second to the third member in this chain of equalities depends on the expression for the interior sum given in [42], p. 137.

For Type IV operators, (A1.8) becomes, by virtue of (3.85),

$$\sum_{j=1}^{m-1} \sum_{k=1}^{n} h^k U_k [(j+1)^{k+1} - 2j^{k+1} + (j-1)^{k+1}]/(k+1)! \tag{A1.13}$$

which readily simplifies to

$$\sum_{k=1}^{n} \frac{U_k h^k}{(k+1)!} [m^{k+1} - (m-1)^{k+1} - 1] \tag{A1.14}$$

or, finally,

$$\sum_{k=1}^{n} \frac{U_k}{(k+1)!} \{ \sum_{j=0}^{k} (-1)^{k-j} (k+1;j) (mh)^j h^{k-j} - h^k \} \tag{A1.15}$$

For Type I operators, it follows from expansion of $P^{-n}(E^{-1})$ in (3.22) that

$$P_j^k = \left(\frac{h}{2}\right)k \sum_{r=0}^{j<k} (k;r)(k+j-r-1;j-r) \tag{A1.16}$$

where the notation indicates that j is the upper limit of the summation for $j \leq k$ and k is the upper limit otherwise.

Hence, using (A1.16), (A1.8) becomes in this case

$$\sum_{j=1}^{m-1} \sum_{k=1}^{n} U_k\{(\frac{h}{2})^k \sum_{r=0}^{j<k} (k;r)(k+j-r-1;j-r)\} \tag{A1.17}$$

or, in a more convenient form,

$$\sum_{k=1}^{n} U_k(\frac{h}{2})^k \sum_{j=1}^{m-1} \sum_{r=0}^{j<k} (k;r)(k+j-r-1;j-r). \tag{A1.18}$$

An expression analogous to that of (3.45) reveals that the interior sum of (A1.18) is a polynomial of degree k-1 in j. Hence, application of our handy device from [42], p. 137, to $\sum_{j=1}^{m-1} j^{k-1}$ evinces the fact that the interior double sum in (A1.18) is a polynomial of degree k in m. As with the preceding two cases, (A1.18) may then be expressed as a polynomial in h with constant coefficients.

BIOGRAPHICAL SKETCH

The author was born on March 12, 1927 at St. Petersburg, Florida, where he received his primary and secondary education. After a brief attendance at Florida Southern College, Lakeland, Florida, he enlisted in the U. S. Navy in 1945 and was honorably discharged after seventeen months of service, much of which was in the Philippines.

All of the remainder of the author's undergraduate and graduate study has been carried out at the University of Florida during two separate periods. During the first of these, which started in September 1947 and ended in February 1952, the author earned the degree Bachelor of Science, with Honors (in July 1950) and completed three semesters of graduate study. His major field in this as well as the subsequent period was mathematics. The second period of the author's attendance at Florida was initiated in September 1955 and has extended to the present date. He received the degree Master of Science in January 1956.

The author has had a varied professional experience; he held a graduate assistantship in the Department of Mathematics from September 1950 to February 1952 and again from September 1955 to June 1956. From February 1952 to August 1955 he was employed as an aeronautical research scientist by the National Advisory Committee for Aeronautics (which has recently become the National Aeronautics and Space Administration) at Langley Field, Virginia. During the summer of 1956 he held the position of physicist at the U. S. Navy Mine Defense Laboratory, Panama City, Florida. From September 1956 to the present, he has been an assistant in research on the faculty of the Department of Engineering Mechanics. As a result of researches carried out in connection with the aforementioned employment, the author has published a number of technical papers.

The author is a member of the honor societies Sigma Xi and Phi Kappa Phi and of the following professional organizations: American Mathematical Society, Society for Industrial and Applied Mathematics and the Institute of Radio Engineers.

This dissertation was prepared under the direction of the chairman of the candidate's supervisory committee and has been approved by all members of the committee. It was submitted to the Dean of the College of Arts and Sciences and to the Graduate Council and was approved as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June 8, 1959

A H Gropp
Dean, College of Arts and Sciences

Dean, Graduate School

SUPERVISORY COMMITTEE:

F W Kokomoor
Chairman

R. C. Meacham
Co-chairman

C. B. Smith

J. T. Moon

R. W. Cowan

William A. Nash